GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 39 Seconds
(without alignments)
1165.089 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHHTLSLPESGPVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1840	99.0	341	21	AAB06206	Immunogenic peptid
2	1595	85.8	341	21	AAB06208	Immunogenic peptid
3	1555	83.7	345	21	AAB06207	Immunogenic peptid
4	1553.5	83.6	342	21	AAB06205	Immunogenic peptid
5	1539	82.8	341	21	AAB06202	Immunogenic peptid
6	1526	82.1	341	21	AAB03644	Opossum IgE heavy
7	1511.5	81.4	342	21	AAB06201	Immunogenic peptid
8	1051	56.6	340	21	AAB03643	Rat IgE heavy chai
9	1012	54.5	313	21	AAY79996	Rat immunoglobulin
10	973	52.4	343	21	AAB06204	Platypus IgE heavy
11	940.5	50.6	421	23	AAU80300	Mouse IgE heavy ch
12	938.5	50.5	426	17	AAR97753	Canine IgE. Canis
13	936	50.4	332	23	AAU80297	Mouse IgE heavy ch
14	936	50.4	332	23	AAU80298	Murine IgE heavy c
15	936	50.4	332	23	AAU80299	Murine IgE heavy c
16	934.5	50.3	417	18	AAW23067	Canine IgE heavy c
17	929.5	50.0	432	23	AAU80294	Murine IgE heavy c
18	925	49.8	343	23	AAU80295	Murine IgE heavy c
19	925	49.8	343	23	AAU80296	Murine IgE heavy c
20	922.5	49.7	312	21	AAY79995	Dog immunoglobulin
21	898	48.3	424	23	AAM50103	Equine IgE heavy c
22	896	48.2	424	23	AAM50104	Equine IgE heavy c
23	862.5	46.4	561	20	AAY17415	Mouse immunoglobul
24	847	45.6	313	21	AAY79997	Mouse immunoglobul
25	745	40.1	323	23	AAU80284	Human IgE heavy ch
26	745	40.1	323	23	AAU80285	Human IgE C2-C3-C4
27	745	40.1	323	23	AAU80286	Human IgE C2-C3-C4
28	745	40.1	324	16	AAR83559	Fc(epsilon) CH2'-C
29	745	40.1	325	16	AAR75225	Human IgE Fc chain
30	745	40.1	325	16	AAR77241	Human IgE Fc chain
31	745	40.1	331	21	AAB03642	Human IgE heavy ch
32	745	40.1	367	9	AAP80291	Interleukin-2/IgE
33	745	40.1	428	23	AAU80283	Human IgE heavy ch
34	745	40.1	428	23	AAM50940	Human IgE epsilon
35	745	40.1	428	23	AAM47863	Human Ig-E heavy c
36	743	40.0	330	23	AAU80289	Human IgE C2-C3-C4
37	742	39.9	325	21	AAY79994	Human immunoglobul
38	741	39.9	325	16	AAR83582	CH2 to CH4 of huma
39	741	39.9	336	23	AAU80288	Human IgE heavy ch
40	741	39.9	441	23	AAU80287	Human IgE heavy ch
41	737	39.7	493	5	AAP40065	Sequence of human
42	726	39.1	315	16	AAR85582	Fc(epsilon) CH2'-C
43	709	38.2	428	14	AAR42950	Human IgE heavy ch
44	525.5	28.3	245	14	AAR38858	Anti-allergic chim
45	520.5	28.0	468	22	AAU69120	Canine IgG heavy c

ALIGNMENTS

RESULT 1
AAB06206
ID AAB06206 standard; protein; 341 AA.
XX
AC AAB06206;
XX
DT 22-NOV-2000 (first entry)
XX
DE Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.
XX
KW Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW asthma; eczema; immunogenic peptide.
XX
OS Chimeric - Didelphis virginiana.
OS Chimeric - Rattus sp.
XX
PN WO200025722-A2.
XX
PD 11-MAY-2000.
XX
PF 21-OCT-1999; 99WO-SE01896.
XX
PR 02-NOV-1998; 98US-0106652.
PR 22-SEP-1999; 99US-0401636.
XX
PA (RESI-) RESISTENTIA PHARM AB.
XX
PI Hellman LT;
XX
DR WPI; 2000-365342/31.
XX
PT Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -

```
XX
PS   Disclosure; Fig 2; 50pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the heavy
CC   chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC   constant region 3 from the rat. It was shown to cause a stronger
CC   polyclonal anti-self IgE response than peptides consisting of the same
CC   regions from one mammal. Immunogenic peptides, particularly those
CC   consisting of different heavy chain constant regions, can be used for
CC   vaccination in humans, against bacterial and viral infections and
CC   allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence   341 AA;

  Query Match             99.0%;  Score 1840;  DB 21;  Length 341;
  Best Local Similarity   99.4%;  Pred. No. 1.8e-144;
  Matches  339;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
              ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSTKHHNATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 2
AAB06208
ID   AAB06208 standard; protein; 341 AA.
XX
AC   AAB06208;
XX
DT   22-NOV-2000  (first entry)
XX
DE   Immunogenic peptide consisting of opossum CH2, dog CH3 and opossum CH4.
XX
KW   Dog; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Chimeric - Didelphis virginiana.
OS   Chimeric - Canis sp.
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SE01896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (RESI-) RESISTENTIA PHARM AB.
XX
PI   Hellman LT;
XX
DR   WPI; 2000-365342/31.
XX
PT   Immunogenic polypeptides useful for preventing the harmful effects of
PT   immunoglobulin E in mammals -
XX
PS   Disclosure; Fig 2; 50pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the heavy
CC   chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC   constant region 3 from the dog. It was shown to cause a stronger
CC   polyclonal anti-self IgE response than peptides consisting of the same
CC   regions from one mammal. Immunogenic peptides, particularly those
CC   consisting of different heavy chain constant regions, can be used for
CC   vaccination in humans, against bacterial and viral infections and
CC   allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence   341 AA;

  Query Match             85.8%;  Score 1595;  DB 21;  Length 341;
  Best Local Similarity   85.0%;  Pred. No. 3.8e-124;
  Matches  290;  Conservative   16;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              :||: :|||| :|| |||||||| :  ||:||||:|| : | : :|| || |: :
Db        121 KCSESDPRGVTSYLSPPSPLDLYVHKAPKITCLVVDLATMEGMNLTWYRESKEPVNPGPL 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
                | : | ::|| |||:  |||||| | ||| ||| || ||||| ||||||||||||
Db        181 NKKDHFNGTITVTSTLPVNTNDWIEGETYYCRVTHPHLPKDIVRSIAKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 3
AAB06207
ID   AAB06207 standard; protein; 345 AA.
XX
AC   AAB06207;
XX
DT   22-NOV-2000  (first entry)
XX
DE   Immunogenic peptide consisting of opossum CH2, pig CH3 and opossum CH4.
XX
KW   Pig; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Chimeric - Didelphis virginiana.
OS   Chimeric - Sus scrofa.
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SE01896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (RESI-) RESISTENTIA PHARM AB.
XX
```

PI Hellman LT;
XX
DR WPI; 2000-365342/31.
XX
PT Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -
XX
PS Disclosure; Fig 2; 50pp; English.
XX
CC The present sequence is an immunogenic peptide consisting of the heavy
CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC constant region 3 from the pig. It was shown to cause a stronger
CC polyclonal anti-self IgE response than peptides consisting of the same
CC regions from one mammal. Immunogenic peptides, particularly those
CC consisting of different heavy chain constant regions, can be used for
CC vaccination in humans, against bacterial and viral infections and
CC allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ Sequence 345 AA;

Query Match 83.7%; Score 1555; DB 21; Length 345;
Best Local Similarity 84.1%; Pred. No. 8e-121;
Matches 290; Conservative 13; Mismatches 38; Indels 4; Gaps 2;

```
Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RC-SDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSI---G 176
              || :: |||||  || ||:||||| : :|||||||:|| | ||: : | || |  :
Db        121 RCTAESEPRGVSAYLSPPTPLDLYVHKSPKLTCLVVDLASSENVNLLWSRENKGGVILPP 180

Qy        177 SASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAP 236
                    : | | || |||:  |||||| | | | ||  ||||:|||:|||||||||
Db        181 PGPPVIKPQFNGTFSATSTLPVNVSDWIEGETYYCNVTHPDLPKPILRSISKLPGKRLAP 240

Qy        237 EVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPS 296
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        241 EVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPS 300

Qy        297 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||||
Db        301 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 345
```

RESULT 4
AAB06205
ID AAB06205 standard; protein; 342 AA.
XX
AC AAB06205;
XX
DT 22-NOV-2000 (first entry)
XX
DE Immunogenic peptide consisting of opossum CH2, human CH3 and opossum CH4.
XX
KW Human; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW asthma; eczema; immunogenic peptide.
XX
OS Chimeric - Didelphis virginiana.
OS Chimeric - Homo sapiens.
XX
PN WO200025722-A2.
XX
PD 11-MAY-2000.
XX
PF 21-OCT-1999; 99WO-SE01896.
XX
PR 02-NOV-1998; 98US-0106652.
PR 22-SEP-1999; 99US-0401636.
XX
PA (RESI-) RESISTENTIA PHARM AB.
XX
PI Hellman LT;
XX
DR WPI; 2000-365342/31.
XX
PT Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -
XX
PS Disclosure; Fig 2; 50pp; English.
XX
CC The present sequence is an immunogenic peptide consisting of the heavy
CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC constant region 3 from the human. It was shown to cause a stronger
CC polyclonal anti-self IgE response than peptides consisting of the same
CC regions from one mammal. Immunogenic peptides, particularly those
CC consisting of different heavy chain constant regions, can be used for
CC vaccination in humans, against bacterial and viral infections and
CC allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ Sequence 342 AA;

Query Match 83.6%; Score 1553.5; DB 21; Length 342;
Best Local Similarity 82.7%; Pred. No. 1.1e-120;
Matches 283; Conservative 21; Mismatches 37; Indels 1; Gaps 1;

```
Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSAS 179
              :|:|  ||||  ||  ||| ||:   :| :||||:||  |:  : :|| |   | :  ::
Db        121 KCADSNPRGVSAYLSRPSPFDLFIRKSPTITCLVVDLAPSKGTVNLTWSRASGKPVNHST 180

Qy        180 QRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
              ::  |  : | ::|| |||  :|||||| ||||| ||| |: ::|| ||||||||||||
Db        181 RKEEKQRNGTLTVTSTLPVGTRDWIEGETYQCRVTHPHLPRALMRSTTKLPGKRLAPEVY 240

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        241 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFL 300

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 342
```

RESULT 5
AAB06202
ID AAB06202 standard; protein; 341 AA.
XX
AC AAB06202;
XX
DT 22-NOV-2000 (first entry)
XX
DE Immunogenic peptide consisting of opossum CH2, mouse CH3 and opossum CH4.
XX
KW Mouse; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW asthma; eczema; immunogenic peptide.
XX
OS Chimeric - Didelphis virginiana.
OS Chimeric - Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..341
FT /label= OTHER
FT /note= "Xaa=unknown"
XX

PN WO200025722-A2.
XX
PD 11-MAY-2000.
XX
PF 21-OCT-1999; 99WO-SE01896.
XX
PR 02-NOV-1998; 98US-0106652.
PR 22-SEP-1999; 99US-0401636.
XX
PA (RESI-) RESISTENTIA PHARM AB.
XX
PI Hellman LT;
XX
DR WPI; 2000-365342/31.
XX
PT Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -
XX
PS Disclosure; Fig 2; 50pp; English.
XX
CC The present sequence is an immunogenic peptide consisting of the heavy
CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC constant region 3 from the mouse. It was shown to cause a stronger
CC polyclonal anti-self IgE response than peptides consisting of the same
CC regions from one mammal. Immunogenic peptides, particularly those
CC consisting of different heavy chain constant regions, can be used for
CC vaccination in humans, against bacterial and viral infections and
CC allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ Sequence 341 AA;

Query Match 82.8%; Score 1539; DB 21; Length 341;
Best Local Similarity 85.9%; Pred. No. 1.7e-119;
Matches 293; Conservative 12; Mismatches 34; Indels 2; Gaps 2;

```
Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||:|||||||||||||||||||||:|||||| ||:|||||||  |||||||||
Db         61 LVDGQEAEHLFPYTTRPKREGGQTFSLQSEVHITQGQWHSSHTYTCHVKXXGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              || | |||||||||||||||||::| ||||||||||||::| ||| :| | :   | |
Db        121 RCPDHEPRGVITYLIPPSPLDLYQHGAPKLTCLVLDLESEKHINVTWNQE-KPTXXXAXQ 179

Qy        181 RSTK-HHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
               || |||||| |||||||  ||||||  ||   ||||  |  | |||:|||||||| ||
Db        180 WYTKHHHHATTXITSILPVVXKDWIEGYXYQXIXDHPHPIKXXVPSITRLPGKRLAPXVY 239

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              |||||||||  ||||||||||||||||||||||   |||||||||||||||||||| ||||
Db        240 MLPPSPEETXTTRTVTCLIRGFYPSEISVQWLFXXEEDHTGHHTTTRPQKDHGTDXSFFL 299

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              ||||||:|||||||:|||||||||||||||||||||||||
Db        300 YSRMLVHKSIWEKGHLVTCRVVHEALPGSRTLEKSLHYSAG 340
```

RESULT 6
AAB03644
ID AAB03644 standard; protein; 341 AA.
XX
AC AAB03644;
XX
DT 22-NOV-2000 (first entry)
XX
DE Opossum IgE heavy chain constant regions 2, 3 and 4.
XX
KW Opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW asthma; eczema; immunogenic peptide.
XX

OS Didelphis virginiana.
XX
PN WO200025722-A2.
XX
PD 11-MAY-2000.
XX
PF 21-OCT-1999; 99WO-SE01896.
XX
PR 02-NOV-1998; 98US-0106652.
PR 22-SEP-1999; 99US-0401636.
XX
PA (RESI-) RESISTENTIA PHARM AB.
XX
PI Hellman LT;
XX
DR WPI; 2000-365342/31.
XX
PT Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -
XX
PS Disclosure; Fig 1; 50pp; English.
XX
CC The present sequence is an immunogenic peptide consisting of the heavy
CC chain constant regions 2, 3 and 4 of the opossum IgE. It was used to
CC construct a number of immunogenic peptides which consisted of regions of
CC IgE from different mammals, which appear to cause a stronger polyclonal
CC anti-self IgE response than peptides consisting of the same regions from
CC one mammal. Immunogenic peptides, particularly those consisting of
CC different heavy chain constant regions, can be used for vaccination
CC in humans, against bacterial and viral infections and allergies, such
CC as asthma, fur, pollen and food allergies and eczema.
XX
SQ Sequence 341 AA;

Query Match 82.1%; Score 1526; DB 21; Length 341;
Best Local Similarity 80.4%; Pred. No. 2e-118;
Matches 274; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

```
Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||::
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVRHNGSIFEDSAQ 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              :||| :|||:  |::||:| ||:    | : ||::|| | ||: ||| ||    :  :|
Db        121 KCSDTDPRGISAYILPPTPQDLFVKKVPTIGCLIVDLASAENVKVTWSRESGGPVNPSSL 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
              : :: | ::|| |||:  |||||: | ||:: |  | |::|:|:| ||||||||||
Db        181 VVKEQYNGTFTVTSHLPVNTDDWIEGDTYTCRLESPDMPVPLIRTISKAPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
```

RESULT 7
AAB06201
ID AAB06201 standard; protein; 342 AA.
XX
AC AAB06201;
XX
DT 22-NOV-2000 (first entry)
XX
DE Immunogenic peptide consisting of opossum CH2, CH4 and rat/opossum CH3.
XX

```
KW   Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Chimeric - Didelphis virginiana.
OS   Chimeric - Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..342
FT                   /label= OTHER
FT                   /note= "Xaa=unknown"
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SE01896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (RESI-) RESISTENTIA PHARM AB.
XX
PI   Hellman LT;
XX
DR   WPI; 2000-365342/31.
XX
PT   Immunogenic polypeptides useful for preventing the harmful effects of
PT   immunoglobulin E in mammals -
XX
PS   Disclosure; Fig 2; 50pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the heavy
CC   chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC   constant region 3 created from a combination of the one from the rat
CC   and the one from the opossum. It was shown to cause a stronger
CC   polyclonal anti-self IgE response than peptides consisting of the same
CC   regions from one mammal. Immunogenic peptides, particularly those
CC   consisting of different heavy chain constant regions, can be used for
CC   vaccination in humans, against bacterial and viral infections and
CC   allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence   342 AA;

  Query Match             81.4%;  Score 1511.5;  DB 21;  Length 342;
  Best Local Similarity   83.6%;  Pred. No. 3.2e-117;
  Matches  285;  Conservative   16;  Mismatches   39;  Indels    1;  Gaps    1;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||:||||||||||||||||||||||:|||||| ||:|||||| |||||||||
Db         61 LVDGQEAEHLFPYTTRPKREGGQTFSLQSEVHITQGQWHSSHTYTCHVKXXGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              |||||||||||||||||||||||:|||||||||||||||:| |||| || | || |
Db        121 RCSDDEPRGVITYLIPPSPLDLYEHGTPKLTCLVLDLESEEHIXVTWVXERPKXIGXARS 180

Qy        181 RSTKH-HHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                  |  :| | : || |||:  |||||  |   :: |     :: :|: |||||||| ||
Db        181 LVVKEXYHGTFTXTSHLPVETDDWIEGXTYTXXLESPDMIVILIPTISALPGKRLAPXVY 240

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              |||||||| |||||||||||||||||||   |||||||||||||||||| ||||
Db        241 MLPPSPEETXTTRTVTCLIRGFYPSEISVQWLFXXEEDHTGHHTTTRPQKDHGTDXSFFL 300

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              ||||||:||||||||:||||||||||||||||||||||||
Db        301 YSRMLVHKSIWEKGHLVTCRVVHEALPGSRTLEKSLHYSAG 341


RESULT 8
AAB03643
ID   AAB03643 standard; protein; 340 AA.
XX
AC   AAB03643;
XX
DT   22-NOV-2000  (first entry)
XX
DE   Rat IgE heavy chain constant regions 2, 3 and 4.
XX
KW   Rat; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Rattus sp.
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SE01896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (RESI-) RESISTENTIA PHARM AB.
XX
PI   Hellman LT;
XX
DR   WPI; 2000-365342/31.
XX
PT   Immunogenic polypeptides useful for preventing the harmful effects of
PT   immunoglobulin E in mammals -
XX
PS   Disclosure; Fig 1; 50pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the
CC   heavy chain constant regions 2, 3 and 4 of the rat IgE. It was used to
CC   construct a number of immunogenic peptides which consisted of regions of
CC   IgE from different mammals, which appear to cause a stronger polyclonal
CC   anti-self IgE response than peptides consisting of the same regions from
CC   one mammal. Immunogenic peptides, particularly those consisting of
CC   different heavy chain constant regions, can be used for vaccination
CC   in humans, against bacterial and viral infections and allergies, such
CC   as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence   340 AA;

  Query Match             56.6%;  Score 1051;  DB 21;  Length 340;
  Best Local Similarity   62.6%;  Pred. No. 5.2e-79;
  Matches  206;  Conservative   35;  Mismatches   78;  Indels   10;  Gaps    3;

Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              || |  ||| | ||||||    ||||||| | | |      | : ||:| ::       | |
Db          8 PVNITKPTVDLLHSSCDPNA-FHSTIQLYCFVYGHIQNDVSIHWLMDDRKI-----YETH 61

Qy         77 PK----REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVIT 132
              :    :| |:  |  | :|||| |||| :|:|| |   |  :  :||||||||||||
Db         62 AQNVLIKEEGKLASTYSRLNITQQQWMSESTFTCKVTSQGENYWAHTRRCSDDEPRGVIT 121

Qy        133 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSI 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        122 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHNATTSI 181

Qy        193 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR 252
              |||||||||||||||||||||||||||||||| |||| |||||:  |  ||     |
Db        182 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKAPGKRSAPEVYVFLPPEEEEKDKR 241

Qy        253 TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEK 312
              |:|||: |:| :|||||| :::       |:|| | | :|::  ||::||: | |::| :
Db        242 TLTCLIQNFFPEDISVQWLQDSKLIPKSQHSTTTPLKYNGSNQRFFIFSRLEVTKALWTQ 301

Qy        313 GNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
                  ||||:||||    | ||:::  | ||
Db        302 TKQFTCRVIHEALREPRKLERTISKSLGN 330
```

```
RESULT 9
AAY79996
ID   AAY79996 standard; Protein; 313 AA.
XX
AC   AAY79996;
XX
DT   15-MAY-2000  (first entry)
XX
DE   Rat immunoglobulin E epsilon heavy chain SEQ ID NO:3.
XX
KW   Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW   immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW   antibody; allergy; allergic disease; immunisation; anti-allergic;
KW   anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS   Rattus sp.
XX
PN   WO9967293-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13959.
XX
PR   20-JUN-1998;   98US-0100287.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Walfield AM;
XX
DR   WPI; 2000-160578/14.
XX
PT   New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT   for immunization against allergy -
XX
PS   Example 1; Page 66-68; 155pp; English.
XX
CC   The present invention describes immunoglobulin E (IgE)-CH3 domain
CC   antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC   and anti-asthmatic properties. (I) induces polyclonal antibodies
CC   specific for a target effector site on the epsilon-heavy chain of IgE,
CC   and so preventing triggering and activation of mast cells and basophils
CC   and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC   containing (I) are used for active immunisation against IgE-mediated
CC   allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC   dermatitis. Nucleic acids that encode these compounds are useful for
CC   recombinant production of corresponding peptides or in DNA vaccines.
CC   Conjugates of (I) that include a promiscuous T helper cell epitope
CC   (functional in genetically diverse subjects), in addition to a B cell
CC   target epitope, have increased immunogenicity and may include cyclic
CC   constraints (disulfide bridge) to stabilise conformational features and
CC   maximize cross-reactivity to the natural target. They induce safe
CC   (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC   acid sequences used in the exemplification of the present invention.
XX
SQ   Sequence   313 AA;

  Query Match             54.5%;  Score 1012;  DB 21;  Length 313;
  Best Local Similarity   63.3%;  Pred. No. 8e-76;
  Matches  198;  Conservative   32;  Mismatches   73;  Indels   10;  Gaps    3;

Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              ||  |   ||| | ||||||    ||||||  | | |      | : ||:| ::       |.|
Db          3 PVNITKPTVDLLHSSCDPNA-FHSTIQLYCFVYGHIQNDVSIHWLMDDRKI-----YDTH 56

Qy         77 PK----REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVIT 132
              :    :| |:  |  | :|||| |||| :|:|| |   |  :   :|||||||||||||
Db         57 AQNVLIKEEGKLASTYSRLNITQQQWMSESTFTCKVTSQGENYWAHTRRCSDDEPRGVIT 116

Qy        133 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSI 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        117 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHNATTSI 176

Qy        193 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR 252
              ||||||||||||||||||||||||||||||||||  | | |||||:  |  ||    |
Db        177 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKALGLRSAPEVYVFLPPEEEEKNKR 236

Qy        253 TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEK 312
              |:|||:|:|: :|||||| :::      |:|| | | :|::  ||::||: | |::| :
Db        237 TLTCLIQNFFPEDISVQWLQDSKLIPKSQHSTTTPLKTNGSNQRFFIFSRLEVTKALWTQ 296

Qy        313 GNLVTCRVVHEAL 325
                  ||||:||||
Db        297 TKQFTCRVIHEAL 309


RESULT 10
AAB06204
ID   AAB06204 standard; protein; 343 AA.
XX
AC   AAB06204;
XX
DT   22-NOV-2000  (first entry)
XX
DE   Platypus IgE heavy chain constant regions 2, 3 and 4.
XX
KW   Platypus; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Ornithorhynchus anatinus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..343
FT                   /label= OTHER
FT                   /note= "Xaa=unknown"
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SE01896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (RESI-) RESISTENTIA PHARM AB.
XX
PI   Hellman LT;
XX
DR   WPI; 2000-365342/31.
XX
PT   Immunogenic polypeptides useful for preventing the harmful effects of
PT   immunoglobulin E in mammals -
XX
PS   Disclosure; Fig 2; 50pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the heavy
CC   chain constant regions 2, 3 and 4 of the platypus IgE. It was used to
CC   construct a number of immunogenic peptides which consisted of regions of
CC   IgE from different mammals, which appear to cause a stronger polyclonal
CC   anti-self IgE response than peptides consisting of the same regions from
CC   one mammal. Immunogenic peptides, particularly those consisting of
CC   different heavy chain constant regions, can be used for vaccination
CC   in humans, against bacterial and viral infections and allergies, such
CC   as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence   343 AA;

  Query Match             52.4%;  Score 973;  DB 21;  Length 343;
  Best Local Similarity   52.3%;  Pred. No. 1.6e-72;
  Matches  180;  Conservative   58;  Mismatches  100;  Indels    6;  Gaps    4;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              |||||||||    :|    .|||||| |||||||||: ::|:||||::|:||| : | |
Db          1 EFHHHHHHTEVYSDSSK-DPIPPTVKLLHSSCDPRGDSQASIELLCLITGYSPAGIQVDW 59
```

```
Qy          61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKH--NGSIFEDS 118
               |||||:||:|||||  |||||  ::||  |||:||| ||:|  |:|| |    :   ::||
Db          60 LVDGQKAEHLFPYTAPPKREGNRSFSSHSEVHITQDQWLSGKTFTCQVTXLADKKTYQDS 119

Qy         119 SRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSA 178
               :|:|:| :|||:  :| |||| ||| : |||||||::|| | | : ||| ||    : :
Db         120 ARKCADSDPRGITVFLTPPSPTDLYISKTPKLTCLIIDLVSTEGMEVTWSRESGTPLXAE 179

Qy         179 SQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEV 238
               |    |  :  | :||  :|| ||| | |   ||    ||::::|||||||||| |
Db         180 SPEEGKXFXGTMTFISTVPVXIQDWNEGESYTCXXAHPDLISPIIKTVTKLPGKRLAPXV 239

Qy         239 YMLPPSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPS 296
               |  ||   |     : ::|||||||||  |||:|| : :   | |: ||:| || | | :
Db         240 YAFPPHQAEVSHXASLSLTCLIRGFYPENISVRWLLDXKPLPTEHYRTTKPLKDQGPDXA 299

Qy         297 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
               :|||||: |:|| ||:|:: ||:|||||| ||  |:   :::|
Db         300 YFLYSRLAVHKSTWEQGHVYTCQVVHEALP-SRNTERKFQHTSG 342
```

RESULT 11
AAU80300
ID AAU80300 standard; Protein; 421 AA.
XX
AC AAU80300;
XX
DT 30-JUL-2002 (first entry)
XX
DE Mouse IgE heavy chain C1-C2-C3-C4 domains.
XX
KW IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;
KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW heavy chain C domain.
XX
OS Mus sp.
XX
PN WO200220038-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-DK00579.
XX
PR 06-SEP-2000; 2000DK-0001326.
PR 15-SEP-2000; 2000US-232831P.
XX
PA (PHAR-) PHARMEXA AS.
XX
PI Klysner S, Von Hoegen P, Voldborg B, Gautam A;
XX
DR WPI; 2002-383033/41.
XX
PT Inducing immune response against autologous immunoglobulin E in an
PT animal, by effecting simultaneous presentation of cytotoxic T
PT lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT -
XX
PS Examples; Page 145-147; 151pp; English.
XX
CC This invention relates to a novel method for inducing an immune response
CC against autologous immunoglobulin E (IgE) in an animal. The method
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC epitope (TH epitope) which is foreign to the animal, by antigen
CC presenting cells (APCs) of the animal's immune system. The epitopes
CC of the invention may be used as a vaccine against allergic diseases. The
CC method of the invention is useful for inducing an immune response
CC against autologous IgE in an animal, which is useful for downregulating
CC autologous IgE in the animal. This method is useful in the prevention
CC and treatment of allergic diseases such as anaphylaxis, allergic
CC rhinitis, asthma and atopic dermatitis. The present sequence represents
CC the mouse IgE heavy chain C1-C2-C3-C4 domains used to create the
CC epitopes of the invention.
XX
SQ Sequence 421 AA;

Query Match 50.6%; Score 940.5; DB 23; Length 421;
Best Local Similarity 55.2%; Pred. No. 1e-69;
Matches 191; Conservative 46; Mismatches 100; Indels 9; Gaps 5;

```
Qy           2 FHHHHHHTLSLPES-----GPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKV 56
               |  |  |  |  ||      || |  ||::| ||||||    |||||| | : |     |
Db          73 FTCHVTHPPSFNESRTILVRPVNITEPTLELLHSSCDPNA-FHSTIQLYCFIYGHILNDV 131

Qy          57 HVTWLVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
               |:||:| :|  :    |  | | |:  |  |::|||: |||| :|:|| |   |  :
Db         132 SVSWLMDDREITDTLAQTVLIKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYL 190

Qy         117 DSSRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIG 176
                 :||| | ||||||||||||||||:|| |||||||:|||||:|: ||| :|:| |:
Db         191 AHTRRCPDHEPRGVITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV- 249

Qy         177 SASQRSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLA 235
               ||||  |||| :|||||||||| ||||||| |||| |||| |||||||||| ||:| |
Db         250 SASQWYTKHHNNATTSITSILPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSA 309

Qy         236 PEVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDP 295
               ||||: ||  ||:   ||:||||: |:| :|||||| : :      |:|| | | :|::
Db         310 PEVYVFPPPEEESEDKRTLTCLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNQ 369

Qy         296 SFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
               ||::||: | |::| :    ||:|:|||   | |||::  | ||
Db         370 GFPIFSRLEVAKTLWTQRKQFTCQVIHEALQKPRKLEKTISTSLGN 415
```

RESULT 12
AAR97753
ID AAR97753 standard; Protein; 426 AA.
XX
AC AAR97753;
XX
DT 28-AUG-1996 (first entry)
XX
DE Canine IgE.
XX
KW IgE; immunoglobulin; antibody; vaccine; allergy; hypersensitivity.
XX
OS Canis familiaris.
XX
PN WO9614867-A1.
XX
PD 23-MAY-1996.
XX
PF 03-NOV-1995; 95WO-US13795.
XX
PR 09-NOV-1994; 94US-0336891.
PR 09-NOV-1994; 94US-0336583.
XX
PA (MERI) MERCK & CO INC.
XX
PI Hollis GF, Patel MD;
XX
DR WPI; 1996-277321/28.
DR N-PSDB; AAT29824.
XX
PT New DNA encoding canine IgE and IgA - useful in vaccines, antisense
PT therapy, assays, drug screening, etc.
XX
PS Claim 11; Page 29-30; 49pp; English.
XX
CC The canine IgE amino acid sequence (AAR97753) was deduced from
CC an isolated gene (AAT29824) obtd. from a canine liver DNA library.
CC The cloning of the IgE gene allows prodn. of large quantities of
CC recombinant IgE using bacterial, yeast, mammalian, insect or

```
CC   viral systems.  The IgE can be used in drug development (e.g.
CC   small molecule screening, assay development and anti-IgE
CC   antibody generation).  Fragments of IgE can be used in vaccines
CC   or to prevent IgE-mediated hypersensitivity.  The new sequence
CC   information permits targeted modulation of IgE-mediated immune
CC   reponses.
XX
SQ   Sequence   426 AA;

  Query Match            50.5%;  Score 938.5;  DB 17;  Length 426;
  Best Local Similarity   53.2%;  Pred. No. 1.5e-69;
  Matches  173;  Conservative   53;  Mismatches   96;  Indels    3;  Gaps    2;

Qy         18 VTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRP 77
              :  |||||||||||||:| || |:|||||||:||: |  : | ||||||:| |:||||
Db        102 LNFIPPTVKLFHSSCNPVGDTHTTIQLLCLISGYVPGDMEVIWLVDGQKATNIFPYTAPG 161

Qy         78 KREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIPP 137
               :||  | |  ||:|||||:|:|  |||| | : |  |:| :|:||: :|||| : | ||
Db        162 TKEGNVT-STHSELNITQGEWVSQKTYTCQVTYQGFTFKDEARKCSESDPRGVTSTLSPP 220

Qy        138 SPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITSILP 197
              ||||||  :  ||:||||:|| : | : :|| || |: :          | : | ::|| ||
Db        221 SPLDLYVHKAPKITCLVVDLATMEGMNLTWYRESKEPVNPGPLNKKDHFNGTITVTSTLP 280

Qy        198 VDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR--TVT 255
              |:  |||||| | ||| ||| || ||||| | ||||  |:||:  |  |: ||    |:|
Db        281 VNTNDWIEGETYYCRVTHPHLPKDIVRSIAKAPGKRAPPDVYLFLPPEEQQGTKDRVTLT 340

Qy        256 CLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNL 315
              |||: |:|::|||||| |:    | :|||  | |  |: |:||::||: |::  ||: |
Db        341 CLIQNFFPADISVQWLRNDSPIQTDQYTTTGPHKVSGSRPAFFIFSRLEVSRVDWEQKNK 400

Qy        316 VTCRVVHEALPGSRTLEKSLHYSAG 340
              ||:|||||| ||| |:| :  : |
Db        401 FTCQVVHEALSGSRILQKWVSKTPG 425


RESULT 13
AAU80297
ID   AAU80297 standard; Protein; 332 AA.
XX
AC   AAU80297;
XX
DT   30-JUL-2002  (first entry)
XX
DE   Mouse IgE heavy chain C2-C3-C4 domain with MIGIS fragment.
XX
KW   IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
KW   antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;
KW   vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW   heavy chain C domain; MIGIS.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO200220038-A2.
XX
PD   14-MAR-2002.
XX
PF   06-SEP-2001; 2001WO-DK00579.
XX
PR   06-SEP-2000; 2000DK-0001326.
PR   15-SEP-2000; 2000US-232831P.
XX
PA   (PHAR-) PHARMEXA AS.
XX
PI   Klysner S,  Von Hoegen P,  Voldborg B,  Gautam A;
XX
DR   WPI; 2002-383033/41.
XX
PT   Inducing immune response against autologous immunoglobulin E in an
PT   animal, by effecting simultaneous presentation of cytotoxic T
PT   lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT   -
XX
PS   Examples; Page 137-138; 151pp; English.
XX
CC   This invention relates to a novel method for inducing an immune response
CC   against autologous immunoglobulin E (IgE) in an animal. The method
CC   comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC   (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC   epitope (TH epitope) which is foreign to the animal, by antigen
CC   presenting cells (APCs) of the animal's immune system. The epitopes
CC   of the invention may be used as a vaccine against allergic diseases. The
CC   method of the invention is useful for inducing an immune response
CC   against autologous IgE in an animal, which is useful for downregulating
CC   autologous IgE in the animal. This method is useful in the prevention
CC   and treatment of allergic diseases such as anaphylaxis, allergic
CC   rhinitis, asthma and atopic dermatitis. The present sequence represents
CC   the mouse IgE heavy chain C2-C3-C4 domain with the MIGIS fragment used
CC   to create the epitopes used in the method of the invention.
XX
SQ   Sequence   332 AA;

  Query Match            50.4%;  Score 936;  DB 23;  Length 332;
  Best Local Similarity   56.7%;  Pred. No. 1.8e-69;
  Matches  185;  Conservative   46;  Mismatches   91;  Indels    4;  Gaps    4;

Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              || |  ||::| ||||||    ||||||  | : |      | |:||:| :|  :     |
Db          4 PVNITEPTLELLHSSCDPNA-FHSTIQLYCFIYGHILNDVSVSWLMDDREITDTLAQTVL 62

Qy         77 PKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIP 136
              | | |:  |  |::|||: |||| :|:|| |    |  :   :||| | ||||||||||
Db         63 IKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYLAHTRRCPDHEPRGVITYLIP 121

Qy        137 PSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHH-HATTSITSI 195
              |||||||:|| |||||||:|||||:|: ||| :|:|  |: ||||  |||| :||||||||
Db        122 PSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV-SASQWYTKHHNNATTSITSI 180

Qy        196 LPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTRTVT 255
              ||| ||||||| |||| |||| |||||||||||| ||:| |||||: ||  ||:   ||:|
Db        181 LPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSAPEVYVFPPPEEESEDKRTLT 240

Qy        256 CLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNL 315
              |||: |:| :|||||| : :       |:|| | | :|::  ||::||: | |::| :
Db        241 CLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNQGFFIFSRLEVAKTLWTQRKQ 300

Qy        316 VTCRVVHEALPGSRTLEKSLHYSAGN 341
              ||:|:||||   | |||::  | ||
Db        301 FTCQVIHEALQKPRKLEKTISTSLGN 326


RESULT 14
AAU80298
ID   AAU80298 standard; Protein; 332 AA.
XX
AC   AAU80298;
XX
DT   30-JUL-2002  (first entry)
XX
DE   Murine IgE heavy chain C2-C3-C4 for mammalian expression.
XX
KW   IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
KW   antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;
KW   vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW   heavy chain C domain.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO200220038-A2.
XX
PD   14-MAR-2002.
XX
```

```
PF   06-SEP-2001; 2001WO-DK00579.
XX
PR   06-SEP-2000; 2000DK-0001326.
PR   15-SEP-2000; 2000US-232831P.
XX
PA   (PHAR-) PHARMEXA AS.
XX
PI   Klysner S,  Von Hoegen P,  Voldborg B,  Gautam A;
XX
DR   WPI; 2002-383033/41.
DR   N-PSDB; ABK51141.
XX
PT   Inducing immune response against autologous immunoglobulin E in an
PT   animal, by effecting simultaneous presentation of cytotoxic T
PT   lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT   -
XX
PS   Disclosure; Page 140-141; 151pp; English.
XX
CC   This invention relates to a novel method for inducing an immune response
CC   against autologous immunoglobulin E (IgE) in an animal. The method
CC   comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC   (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC   epitope (TH epitope) which is foreign to the animal, by antigen
CC   presenting cells (APCs) of the animal's immune system. The epitopes
CC   of the invention may be used as a vaccine against allergic diseases. The
CC   method of the invention is useful for inducing an immune response
CC   against autologous IgE in an animal, which is useful for downregulating
CC   autologous IgE in the animal. This method is useful in the prevention
CC   and treatment of allergic diseases such as anaphylaxis, allergic
CC   rhinitis, asthma and atopic dermatitis. The present sequence represents
CC   he mouse IgE heavy chain C2-C3-C4 domain optimised for a mammalian
CC   expression system used to create the epitopes used in the method of the
CC   invention.
XX
SQ   Sequence   332 AA;

  Query Match             50.4%;  Score 936;  DB 23;  Length 332;
  Best Local Similarity   56.7%;  Pred. No. 1.8e-69;
  Matches  185;  Conservative   46;  Mismatches   91;  Indels    4;  Gaps    4;

Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              || |  ||::| ||||||    |||||| | : |      | |:||:| :|  :    |
Db          4 PVNITEPTLELLHSSCDPNA-FHSTIQLYCFIYGHILNDVSVSWLMDDREITDTLAQTVL 62

Qy         77 PKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIP 136
              | | |:  |  |::||| |||| :|:|| |   |  :   :||| | ||||||||||
Db         63 IKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYLAHTRRCPDHEPRGVITYLIP 121

Qy        137 PSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHH-HATTSITSI 195
              |||||||:|| |||||||:|||||:|: ||| :|:| |: |||| |||| :||||||||
Db        122 PSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV-SASQWYTKHHNNATTSITSI 180

Qy        196 LPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTRTVT 255
              ||| ||||||| |||| |||| |||||||||||| ||:| |||||: ||  ||:   ||:|
Db        181 LPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSAPEVYVFPPPEEESEDKRTLT 240

Qy        256 CLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNL 315
              |||: |:| :|||||| : :      |:|| | | :|::  ||::||: | |::| :
Db        241 CLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNQGFFIFSRLEVAKTLWTQRKQ 300

Qy        316 VTCRVVHEALPGSRTLEKSLHYSAGN 341
              ||:|:||||   | |||::  | ||
Db        301 FTCQVIHEALQKPRKLEKTISTSLGN 326


RESULT 15
AAU80299
ID   AAU80299 standard; Protein; 332 AA.
XX
AC   AAU80299;
XX
DT   30-JUL-2002  (first entry)
XX
DE   Murine IgE heavy chain C2-C3-C4 for E.Coli expression.
XX
KW   IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
KW   antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;
KW   vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW   heavy chain C domain.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO200220038-A2.
XX
PD   14-MAR-2002.
XX
PF   06-SEP-2001; 2001WO-DK00579.
XX
PR   06-SEP-2000; 2000DK-0001326.
PR   15-SEP-2000; 2000US-232831P.
XX
PA   (PHAR-) PHARMEXA AS.
XX
PI   Klysner S,  Von Hoegen P,  Voldborg B,  Gautam A;
XX
DR   WPI; 2002-383033/41.
DR   N-PSDB; ABK51142.
XX
PT   Inducing immune response against autologous immunoglobulin E in an
PT   animal, by effecting simultaneous presentation of cytotoxic T
PT   lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT   -
XX
PS   Disclosure; Page 144-145; 151pp; English.
XX
CC   This invention relates to a novel method for inducing an immune response
CC   against autologous immunoglobulin E (IgE) in an animal. The method
CC   comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC   (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC   epitope (TH epitope) which is foreign to the animal, by antigen
CC   presenting cells (APCs) of the animal's immune system. The epitopes
CC   of the invention may be used as a vaccine against allergic diseases. The
CC   method of the invention is useful for inducing an immune response
CC   against autologous IgE in an animal, which is useful for downregulating
CC   autologous IgE in the animal. This method is useful in the prevention
CC   and treatment of allergic diseases such as anaphylaxis, allergic
CC   rhinitis, asthma and atopic dermatitis. The present sequence represents
CC   the mouse IgE heavy chain C2-C3-C4 domain optimised for an E.Coli
CC   expression system used to create the epitopes used in the method of the
CC   invention.
XX
SQ   Sequence   332 AA;

  Query Match             50.4%;  Score 936;  DB 23;  Length 332;
  Best Local Similarity   56.7%;  Pred. No. 1.8e-69;
  Matches  185;  Conservative   46;  Mismatches   91;  Indels    4;  Gaps    4;

Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              || |  ||::| ||||||    |||||| | : |      | |:||:| :|  :    |
Db          4 PVNITEPTLELLHSSCDPNA-FHSTIQLYCFIYGHILNDVSVSWLMDDREITDTLAQTVL 62

Qy         77 PKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIP 136
              | | |:  |  |::||| |||| :|:|| |   |  :   :||| | ||||||||||
Db         63 IKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYLAHTRRCPDHEPRGVITYLIP 121

Qy        137 PSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHH-HATTSITSI 195
              |||||||:|| |||||||:|||||:|: ||| :|:| |: |||| |||| :||||||||
Db        122 PSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV-SASQWYTKHHNNATTSITSI 180

Qy        196 LPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTRTVT 255
              ||| ||||||| |||| |||| |||||||||||| ||:| |||||: ||  ||:   ||:|
Db        181 LPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSAPEVYVFPPPEEESEDKRTLT 240

Qy        256 CLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNL 315
```

```
            |||: |:| :|||||| : :      |:|| | | :|::  ||::||: | |::| :
Db        241 CLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNQGFFIFSRLEVAKTLWTQRKQ 300

Qy        316 VTCRVVHEALPGSRTLEKSLHYSAGN 341
              ||:|:||||   | |||::  | ||
Db        301 FTCQVIHEALQKPRKLEKTISTSLGN 326
```

Search completed: July 9, 2003, 13:41:48
Job time : 41 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:40:31 ; Search time 53 Seconds
(without alignments)
749.040 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHHTLSLPESGPVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1858	100.0	341	9	US-10-176-664-4	Sequence 4, Appli
2	1858	100.0	341	10	US-09-401-636-4	Sequence 4, Appli
3	1840	99.0	341	9	US-10-176-664-9	Sequence 9, Appli
4	1840	99.0	341	10	US-09-401-636-9	Sequence 9, Appli
5	1720	92.6	341	9	US-10-176-664-6	Sequence 6, Appli
6	1720	92.6	341	10	US-09-401-636-6	Sequence 6, Appli
7	1691.5	91.0	342	9	US-10-176-664-5	Sequence 5, Appli
8	1691.5	91.0	342	10	US-09-401-636-5	Sequence 5, Appli
9	1595	85.8	341	9	US-10-176-664-11	Sequence 11, Appl
10	1595	85.8	341	10	US-09-401-636-11	Sequence 11, Appl
11	1555	83.7	345	9	US-10-176-664-10	Sequence 10, Appl
12	1555	83.7	345	10	US-09-401-636-10	Sequence 10, Appl
13	1553.5	83.6	342	9	US-10-176-664-8	Sequence 8, Appli
14	1553.5	83.6	342	10	US-09-401-636-8	Sequence 8, Appli
15	1529	82.3	341	9	US-10-176-664-3	Sequence 3, Appli
16	1529	82.3	341	10	US-09-401-636-3	Sequence 3, Appli
17	1465	78.8	446	9	US-10-214-524-32	Sequence 32, Appl
18	1254	67.5	427	9	US-10-214-524-36	Sequence 36, Appl
19	1051	56.6	340	9	US-10-176-664-2	Sequence 2, Appli
20	1051	56.6	340	10	US-09-401-636-2	Sequence 2, Appli
21	1051	56.6	428	9	US-10-214-524-34	Sequence 34, Appl
22	1025	55.2	343	9	US-10-176-664-7	Sequence 7, Appli
23	1025	55.2	343	10	US-09-401-636-7	Sequence 7, Appli
24	1020.5	54.9	346	9	US-10-152-190-14	Sequence 14, Appl
25	1018.5	54.8	577	9	US-10-214-524-29	Sequence 29, Appl
26	950.5	51.2	426	9	US-10-214-524-28	Sequence 28, Appl
27	940.5	50.6	421	9	US-09-949-375A-28	Sequence 28, Appl
28	940.5	50.6	421	9	US-10-214-524-31	Sequence 31, Appl
29	937.5	50.5	496	9	US-10-214-524-25	Sequence 25, Appl
30	936	50.4	332	9	US-09-949-375A-23	Sequence 23, Appl
31	936	50.4	332	9	US-09-949-375A-25	Sequence 25, Appl
32	936	50.4	332	9	US-09-949-375A-27	Sequence 27, Appl
33	931.5	50.1	431	9	US-09-479-614-14	Sequence 14, Appl
34	931.5	50.1	496	9	US-09-479-614-2	Sequence 2, Appli
35	931.5	50.1	496	9	US-09-479-614-29	Sequence 29, Appl
36	929.5	50.0	432	9	US-09-949-375A-19	Sequence 19, Appl
37	925	49.8	343	9	US-09-949-375A-20	Sequence 20, Appl
38	925	49.8	343	9	US-09-949-375A-22	Sequence 22, Appl
39	865.5	46.6	569	9	US-10-214-524-30	Sequence 30, Appl
40	853	45.9	567	9	US-10-214-524-33	Sequence 33, Appl
41	835	44.9	563	9	US-10-214-524-35	Sequence 35, Appl
42	785	42.2	426	9	US-10-214-524-27	Sequence 27, Appl
43	774.5	41.7	346	9	US-10-152-190-10	Sequence 10, Appl
44	759	40.9	347	9	US-10-152-190-12	Sequence 12, Appl
45	752	40.5	569	9	US-09-847-208-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-176-664-4
; Sequence 4, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-4

Query Match 100.0%; Score 1858; DB 9; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.9e-122;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 2
US-09-401-636-4
; Sequence 4, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellman, Lars T.
;  TITLE OF INVENTION:  ENHANCED VACCINES
;  FILE REFERENCE:  10223/006001
;  CURRENT APPLICATION NUMBER: US/09/401,636
;  CURRENT FILING DATE:  1999-09-22
;  PRIOR APPLICATION NUMBER:  US 60/106,652
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-4

  Query Match            100.0%;  Score 1858;  DB 10;  Length 341;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-122;
  Matches  341;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 3
US-10-176-664-9
; Sequence 9, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
;  APPLICANT: Hellman, Lars T.
;  TITLE OF INVENTION: ENHANCED VACCINES
;  FILE REFERENCE: 10223/006001
;  CURRENT APPLICATION NUMBER: US/10/176,664
;  CURRENT FILING DATE:  2002-06-19
;  PRIOR APPLICATION NUMBER: US/09/401,636
;  PRIOR FILING DATE: 1999-09-22
;  PRIOR APPLICATION NUMBER: US 60/106,652
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-9

  Query Match            99.0%;  Score 1840;  DB 9;  Length 341;
  Best Local Similarity   99.4%;  Pred. No. 8.8e-121;
  Matches  339;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
              ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSTKHHNATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 4
US-09-401-636-9
; Sequence 9, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellman, Lars T.
;  TITLE OF INVENTION:  ENHANCED VACCINES
;  FILE REFERENCE:  10223/006001
;  CURRENT APPLICATION NUMBER: US/09/401,636
;  CURRENT FILING DATE:  1999-09-22
;  PRIOR APPLICATION NUMBER:  US 60/106,652
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-9

  Query Match            99.0%;  Score 1840;  DB 10;  Length 341;
  Best Local Similarity   99.4%;  Pred. No. 8.8e-121;
  Matches  339;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy         121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180

Qy         181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
                ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 RSTKHHNATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240

Qy         241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy         301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
                |||||||||||||||||||||||||||||||||||||||||
Db         301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
```

RESULT 5
US-10-176-664-6
; Sequence 6, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-6

Query Match 92.6%; Score 1720; DB 9; Length 341;
Best Local Similarity 93.3%; Pred. No. 2e-112;
Matches 319; Conservative 9; Mismatches 12; Indels 2; Gaps 2;

```
Qy           1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy          61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy         121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
               || | ||||||||||||||||||:|| |||||||:|||||:|: ||| :|:| |: :|||
Db         121 RCPDHEPRGVITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV-NASQ 179

Qy         181 RSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                 |||| :|||||||||||| ||||||| |||| |||| ||||||||||||||||||||||
Db         180 WYTKHHNNATTSITSILPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKLPGKRLAPEVY 239

Qy         240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299

Qy         300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
               |||||||||||||||||||||||||||||||||||||||||
Db         300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
```

RESULT 6
US-09-401-636-6
; Sequence 6, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-6

Query Match 92.6%; Score 1720; DB 10; Length 341;
Best Local Similarity 93.3%; Pred. No. 2e-112;
Matches 319; Conservative 9; Mismatches 12; Indels 2; Gaps 2;

```
Qy           1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy          61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy         121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
               || | ||||||||||||||||||:|| |||||||:|||||:|: ||| :|:| |: :|||
Db         121 RCPDHEPRGVITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV-NASQ 179

Qy         181 RSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                 |||| :|||||||||||| ||||||| |||| |||| ||||||||||||||||||||||
Db         180 WYTKHHNNATTSITSILPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKLPGKRLAPEVY 239

Qy         240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299

Qy         300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
               |||||||||||||||||||||||||||||||||||||||||
Db         300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
```

RESULT 7
US-10-176-664-5
; Sequence 5, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins

```
US-10-176-664-5

  Query Match             91.0%;  Score 1691.5;  DB 9;  Length 342;
  Best Local Similarity   90.9%;  Pred. No. 2e-110;
  Matches  311;  Conservative   12;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSARS 180

Qy        181 RSTK-HHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                  |  :: | ::|| |||:  |||||: | ||:: |  | |::|:|:| ||||||||||
Db        181 LVVKEQYNGTFTVTSHLPVNTDDWIEGDTYTCRLESPDMPYPLIRTISKAPGKRLAPEVY 240

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 300

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 342


RESULT 8
US-09-401-636-5
; Sequence 5, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellman, Lars T.
;  TITLE OF INVENTION:  ENHANCED VACCINES
;  FILE REFERENCE:  10223/006001
;  CURRENT APPLICATION NUMBER: US/09/401,636
;  CURRENT FILING DATE:  1999-09-22
;  PRIOR APPLICATION NUMBER:  US 60/106,652
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 342
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-5

  Query Match             91.0%;  Score 1691.5;  DB 10;  Length 342;
  Best Local Similarity   90.9%;  Pred. No. 2e-110;
  Matches  311;  Conservative   12;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSARS 180

Qy        181 RSTK-HHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                  |  :: | ::|| |||:  |||||: | ||:: |  | |::|:|:| ||||||||||
Db        181 LVVKEQYNGTFTVTSHLPVNTDDWIEGDTYTCRLESPDMPYPLIRTISKAPGKRLAPEVY 240

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 300

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 342


RESULT 9
US-10-176-664-11
; Sequence 11, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
;  APPLICANT: Hellman, Lars T.
;  TITLE OF INVENTION: ENHANCED VACCINES
;  FILE REFERENCE: 10223/006001
;  CURRENT APPLICATION NUMBER: US/10/176,664
;  CURRENT FILING DATE:  2002-06-19
;  PRIOR APPLICATION NUMBER: US/09/401,636
;  PRIOR FILING DATE: 1999-09-22
;  PRIOR APPLICATION NUMBER: US 60/106,652
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-11

  Query Match             85.8%;  Score 1595;  DB 9;  Length 341;
  Best Local Similarity   85.0%;  Pred. No. 1e-103;
  Matches  290;  Conservative   16;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              :||: :|||| :|| ||||||||  :  ||:||||:|| : | : :|| || |: :
Db        121 KCSESDPRGVTSYLSPPSPLDLYVHKAPKITCLVVDLATMEGMNLTWYRESKEPVNPGPL 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
                  | : | ::|| |||:  |||||| | ||| ||| || ||||| ||||||||||||
Db        181 NKKDHFNGTITVTSTLPVNTNDWIEGETYYCRVTHPHLPKDIVRSIAKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 10
US-09-401-636-11
; Sequence 11, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellman, Lars T.
;  TITLE OF INVENTION:  ENHANCED VACCINES
;  FILE REFERENCE:  10223/006001
;  CURRENT APPLICATION NUMBER: US/09/401,636
;  CURRENT FILING DATE:  1999-09-22
;  PRIOR APPLICATION NUMBER:  US 60/106,652
```

```
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-11

  Query Match             85.8%;  Score 1595;  DB 10;  Length 341;
  Best Local Similarity   85.0%;  Pred. No. 1e-103;
  Matches  290;  Conservative   16;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              :||: :|||| :|| |||||||| :  ||:||||:|| : | : :|| || |: :
Db        121 KCSESDPRGVTSYLSPPSPLDLYVHKAPKITCLVVDLATMEGMNLTWYRESKEPVNPGPL 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
              | : | ::|| |||:  |||||| | ||| ||| || ||||| ||||||||||||
Db        181 NKKDHFNGTITVTSTLPVNTNDWIEGETYYCRVTHPHLPKDIVRSIAKLPGKRLAPEVYM 240

Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341


RESULT 11
US-10-176-664-10
; Sequence 10, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
;  APPLICANT: Hellman, Lars T.
;  TITLE OF INVENTION: ENHANCED VACCINES
;  FILE REFERENCE: 10223/006001
;  CURRENT APPLICATION NUMBER: US/10/176,664
;  CURRENT FILING DATE:  2002-06-19
;  PRIOR APPLICATION NUMBER: US/09/401,636
;  PRIOR FILING DATE: 1999-09-22
;  PRIOR APPLICATION NUMBER: US 60/106,652
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-10

  Query Match             83.7%;  Score 1555;  DB 9;  Length 345;
  Best Local Similarity   84.1%;  Pred. No. 6.5e-101;
  Matches  290;  Conservative   13;  Mismatches   38;  Indels    4;  Gaps    2;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RC-SDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSI---G 176
              || :: |||||  || ||:||||| : :|||||||:|| | ||: : | || |  :
Db        121 RCTAESEPRGVSAYLSPPTPLDLYVHKSPKLTCLVVDLASSENVNLLWSRENKGGVILPP 180

Qy        177 SASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAP 236
                      : | | || |||:  |||||| | | | ||  ||||:|||:|||||||||
Db        181 PGPPVIKPQFNGTFSATSTLPVNVSDWIEGETYYCNVTHPDLPKPILRSISKLPGKRLAP 240

Qy        237 EVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPS 296
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        241 EVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPS 300

Qy        297 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||||
Db        301 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 345


RESULT 12
US-09-401-636-10
; Sequence 10, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellman, Lars T.
;  TITLE OF INVENTION:  ENHANCED VACCINES
;  FILE REFERENCE:  10223/006001
;  CURRENT APPLICATION NUMBER: US/09/401,636
;  CURRENT FILING DATE:  1999-09-22
;  PRIOR APPLICATION NUMBER:  US 60/106,652
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-10

  Query Match             83.7%;  Score 1555;  DB 10;  Length 345;
  Best Local Similarity   84.1%;  Pred. No. 6.5e-101;
  Matches  290;  Conservative   13;  Mismatches   38;  Indels    4;  Gaps    2;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RC-SDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSI---G 176
              || :: |||||  || ||:||||| : :|||||||:|| | ||: : | || |  :
Db        121 RCTAESEPRGVSAYLSPPTPLDLYVHKSPKLTCLVVDLASSENVNLLWSRENKGGVILPP 180

Qy        177 SASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAP 236
                      : | | || |||:  |||||| | | | ||  ||||:|||:|||||||||
Db        181 PGPPVIKPQFNGTFSATSTLPVNVSDWIEGETYYCNVTHPDLPKPILRSISKLPGKRLAP 240

Qy        237 EVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPS 296
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        241 EVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPS 300

Qy        297 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||||
Db        301 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 345


RESULT 13
US-10-176-664-8
```

```
; Sequence 8, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
;  APPLICANT: Hellman, Lars T.
;  TITLE OF INVENTION: ENHANCED VACCINES
;  FILE REFERENCE: 10223/006001
;  CURRENT APPLICATION NUMBER: US/10/176,664
;  CURRENT FILING DATE:  2002-06-19
;  PRIOR APPLICATION NUMBER: US/09/401,636
;  PRIOR FILING DATE: 1999-09-22
;  PRIOR APPLICATION NUMBER: US 60/106,652
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 342
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-8

  Query Match             83.6%;  Score 1553.5;  DB 9;  Length 342;
  Best Local Similarity   82.7%;  Pred. No. 8.1e-101;
  Matches  283;  Conservative   21;  Mismatches   37;  Indels    1;  Gaps    1;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSAS 179
              :|:|  ||||  ||  ||| ||:   :| :||||:|| |:  : :|| |   | :  ::
Db        121 KCADSNPRGVSAYLSRPSPFDLFIRKSPTITCLVVDLAPSKGTVNLTWSRASGKPVNHST 180

Qy        180 QRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
              ::  |  : | ::|| |||  :|||||| ||||| ||| |: ::|| ||||||||||||
Db        181 RKEEKQRNGTLTVTSTLPVGTRDWIEGETYQCRVTHPHLPRALMRSTTKLPGKRLAPEVY 240

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        241 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFL 300

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 342


RESULT 14
US-09-401-636-8
; Sequence 8, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellman, Lars T.
;  TITLE OF INVENTION:  ENHANCED VACCINES
;  FILE REFERENCE:  10223/006001
;  CURRENT APPLICATION NUMBER: US/09/401,636
;  CURRENT FILING DATE:  1999-09-22
;  PRIOR APPLICATION NUMBER:  US 60/106,652
;  PRIOR FILING DATE:  1998-11-02
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 342
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-8

  Query Match             83.6%;  Score 1553.5;  DB 10;  Length 342;
  Best Local Similarity   82.7%;  Pred. No. 8.1e-101;
  Matches  283;  Conservative   21;  Mismatches   37;  Indels    1;  Gaps    1;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSAS 179
              :|:|  ||||  ||  ||| ||:   :| :||||:|| |:  : :|| |   | :  ::
Db        121 KCADSNPRGVSAYLSRPSPFDLFIRKSPTITCLVVDLAPSKGTVNLTWSRASGKPVNHST 180

Qy        180 QRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
              ::  |  : | ::|| |||  :|||||| ||||| ||| |: ::|| ||||||||||||
Db        181 RKEEKQRNGTLTVTSTLPVGTRDWIEGETYQCRVTHPHLPRALMRSTTKLPGKRLAPEVY 240

Qy        240 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFL 299
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        241 MLPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFL 300

Qy        300 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 YSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 342


RESULT 15
US-10-176-664-3
; Sequence 3, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
;  APPLICANT: Hellman, Lars T.
;  TITLE OF INVENTION: ENHANCED VACCINES
;  FILE REFERENCE: 10223/006001
;  CURRENT APPLICATION NUMBER: US/10/176,664
;  CURRENT FILING DATE:  2002-06-19
;  PRIOR APPLICATION NUMBER: US/09/401,636
;  PRIOR FILING DATE: 1999-09-22
;  PRIOR APPLICATION NUMBER: US 60/106,652
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-3

  Query Match             82.3%;  Score 1529;  DB 9;  Length 341;
  Best Local Similarity   80.6%;  Pred. No. 4.1e-99;
  Matches  275;  Conservative   28;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 EFHHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW 60

Qy         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db         61 LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSAQ 120

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              :||| :|||:  |::||:|  ||:   | : ||::|| | ||: ||| ||    :  :|
Db        121 KCSDTDPRGISAYILPPTPQDLFVKKVPTIGCLIVDLASAENVKVTWSRESGGPVNPSSL 180

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYM 240
                 : :: | ::|| |||:  |||||: | ||:: |   | |::|:|:| ||||||||||
Db        181 VVKEQYNGTFTVTSHLPVNTDDWIEGDTYTCRLESPDMPVPLIRTISKAPGKRLAPEVYM 240
```

```
Qy        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        241 LPPSPEETGTTRTVTCLIRGFYPSEISVQWLFNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              |||||||||||||||||||||||||||||||||||||||||
Db        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
```

Search completed: July 9, 2003, 13:50:01
Job time : 54 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 32 Seconds
(without alignments)
2195.690 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHHTLSLPESGPVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	488.5	26.3	375	4	Q9BSZ1	Q9bsz1 homo sapien
2	488.5	26.3	588	4	Q8WUX4	Q8wux4 homo sapien
3	488.5	26.3	597	4	Q9BQB8	Q9bqb8 homo sapien
4	488.5	26.3	597	4	Q96BB9	Q96bb9 homo sapien
5	488.5	26.3	597	4	Q9BU10	Q9bu10 homo sapien
6	488.5	26.3	613	4	Q96EY0	Q96ey0 homo sapien
7	488.5	26.3	613	4	Q8WUK1	Q8wuk1 homo sapien
8	488.5	26.3	614	4	Q96GA6	Q96ga6 homo sapien
9	488.5	26.3	618	4	Q96AA6	Q96aa6 homo sapien
10	463	24.9	337	6	Q95M34	Q95m34 equus cabal
11	455.5	24.5	613	11	Q8VCX7	Q8vcx7 mus musculu
12	455	24.5	473	4	Q8TC63	Q8tc63 homo sapien
13	441.5	23.8	471	4	Q8TC77	Q8tc77 homo sapien
14	435	23.4	437	11	Q9R1A4	Q9r1a4 mus musculu
15	434.5	23.4	463	11	Q99LC4	Q99lc4 mus musculu
16	429.5	23.1	469	11	Q8R3V9	Q8r3v9 mus musculu
17	424.5	22.8	473	11	Q9D8L4	Q9d814 mus musculu
18	406	21.9	473	11	Q91Z05	Q91z05 mus musculu
19	406	21.9	474	11	Q8R3H6	Q8r3h6 mus musculu
20	400.5	21.6	468	11	Q99L31	Q99l31 mus musculu
21	400.5	21.6	473	11	Q99L25	Q99l25 mus musculu
22	390	21.0	701	4	Q96PQ8	Q96pq8 homo sapien
23	355	19.1	416	4	Q9NPP6	Q9npp6 homo sapien
24	347.5	18.7	384	4	Q9UP60	Q9up60 homo sapien
25	346.5	18.6	494	4	Q96K68	Q96k68 homo sapien
26	346.5	18.6	496	4	Q96KX8	Q96kx8 homo sapien
27	345.5	18.6	496	4	Q96DK0	Q96dk0 homo sapien
28	344.5	18.5	497	4	Q8WY24	Q8wy24 homo sapien
29	336.5	18.1	500	4	Q9BRV0	Q9brv0 homo sapien
30	333.5	17.9	426	11	Q9DCD9	Q9dcd9 mus musculu
31	330.5	17.8	486	11	Q91Z07	Q91z07 mus musculu
32	330.5	17.8	487	11	Q99KA4	Q99ka4 mus musculu
33	326	17.5	479	11	Q99M22	Q99m22 mus musculu
34	326	17.5	481	11	Q8VCV5	Q8vcv5 mus musculu
35	326	17.5	484	11	Q99LA6	Q99la6 mus musculu
36	322.5	17.4	684	13	Q90544	Q90544 ginglymosto
37	314	16.9	488	11	Q91WR1	Q91wr1 mus musculu
38	314	16.9	489	11	Q8VCX4	Q8vcx4 mus musculu
39	313	16.8	481	11	Q91WT3	Q91wt3 mus musculu
40	313	16.8	481	11	Q91WT1	Q91wt1 mus musculu
41	313	16.8	482	11	Q91X92	Q91x92 mus musculu
42	313	16.8	484	11	Q8VEA0	Q8vea0 mus musculu
43	312	16.8	480	11	Q91XE1	Q91xe1 mus musculu
44	311	16.7	479	11	Q91WP5	Q91wp5 mus musculu
45	210	11.3	573	4	Q8WU38	Q8wu38 homo sapien

ALIGNMENTS

RESULT 1
Q9BSZ1
ID Q9BSZ1 PRELIMINARY; PRT; 375 AA.
AC Q9BSZ1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 41.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC004476; AAH04476.1; -.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 375 AA; 41314 MW; B1A0A0998F473619 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 375;
Best Local Similarity 31.2%; Pred. No. 2.2e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

```
Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|::  :
Db         24 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 79

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||       |  |  |: : | : | :  |:| : :|| | | |  | :::|  |
```

```
Db         80 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 137

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
              |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        138 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 197

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        198 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 257

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :      :| :
Db        258 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 317

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        318 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 354


RESULT 2
Q8WUX4
ID   Q8WUX4      PRELIMINARY;      PRT;   588 AA.
AC   Q8WUX4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 64.4 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPH;
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC019235; AAH19235.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein.
SQ   SEQUENCE   588 AA;  64438 MW;  FC60DBAD82B39FD7 CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 588;
  Best Local Similarity   31.2%;  Pred. No. 4.1e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|:: :
Db        246 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 301

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                ||        |  |  |: : | : |  :|: : :|| | | |  | :::|  |
Db        302 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 359

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
              |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        360 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 419

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 479

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :      :| :
Db        480 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 539

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        540 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 576


RESULT 3
Q9BQB8
ID   Q9BQB8      PRELIMINARY;      PRT;   597 AA.
AC   Q9BQB8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Unknown (protein for MGC:1905) (protein for MGC:1228).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MUSCLE;
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPH;
RA   Strausberg R.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC006180; AAH06180.1; -.
DR   EMBL; BC001872; AAH01872.1; -.
DR   HSSP; P01825; 7FAB.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 4.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00410; IG_like; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
SQ   SEQUENCE   597 AA;  65300 MW;  2DAFAD50A6375851 CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 597;
  Best Local Similarity   31.2%;  Pred. No. 4.2e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|:: :
Db        246 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 301

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                ||        |  |  |: : | : |  :|: : :|| | | |  | :::|  |
Db        302 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 359

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
              |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        360 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 419

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 479

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :      :| :
Db        480 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 539

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        540 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 576


RESULT 4
```

```
Q96BB9
ID   Q96BB9      PRELIMINARY;      PRT;   597 AA.
AC   Q96BB9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical 65.0 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015760; AAH15760.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 5.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein.
SQ   SEQUENCE   597 AA;  65039 MW;  4FCA3AD8ECE263D9 CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 597;
  Best Local Similarity   31.2%;  Pred. No. 4.2e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy          17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
               |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|::  :
Db         246 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 301

Qy          70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||        |  |  |: : | : | :  |:| : :|| | | |  | :::|  |
Db         302 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 359

Qy         124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
               |:   :  : |||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db         360 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 419

Qy         184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
               | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db         420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 479

Qy         243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
               |: |:     : |:|||: || |::: |||:   :      : |: | :      :| :
Db         480 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 539

Qy         301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
               | : |::  |  |   || | |||||     ||::||
Db         540 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 576


RESULT 5
Q9BU10
ID   Q9BU10      PRELIMINARY;      PRT;   597 AA.
AC   Q9BU10;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 65.3 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPH;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC002963; AAH02963.1; -.
DR   HSSP; P01825; 7FAB.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 4.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00410; IG_like; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein.
SQ   SEQUENCE   597 AA;  65274 MW;  2DAFA8FB7E055851 CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 597;
  Best Local Similarity   31.2%;  Pred. No. 4.2e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy          17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
               |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|::  :
Db         246 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 301

Qy          70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||        |  |  |: : | : | :  |:| : :|| | | |  | :::|  |
Db         302 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 359

Qy         124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
               |:   :  : |||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db         360 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 419

Qy         184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
               | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db         420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 479

Qy         243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
               |: |:     : |:|||: || |::: |||:   :      : |: | :      :| :
Db         480 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 539

Qy         301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
               | : |::  |  |   || | |||||     ||::||
Db         540 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 576


RESULT 6
Q96EY0
ID   Q96EY0      PRELIMINARY;      PRT;   613 AA.
AC   Q96EY0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Unknown (protein for MGC:20337).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011857; AAH11857.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00408; IGc2; 2.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Immunoglobulin domain.
SQ   SEQUENCE   613 AA;  67273 MW;  31214203FB8421E7 CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 613;
  Best Local Similarity   31.2%;  Pred. No. 4.3e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy          17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
               |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|::  :
```

```
Db        241 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 296

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123

Db        297 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 354

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183

Db        355 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 414

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242

Db        415 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 474

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Db        475 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 534

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334

Db        535 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 571


RESULT 7
Q8WUK1
ID   Q8WUK1      PRELIMINARY;      PRT;   613 AA.
AC   Q8WUK1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 67.3 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TONSIL;
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC020240; AAH20240.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein.
SQ   SEQUENCE   613 AA;  67296 MW;  60C7F5950671E315 CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 613;
  Best Local Similarity   31.2%;  Pred. No. 4.3e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69

Db        241 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 296

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123

Db        297 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 354

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183

Db        355 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 414

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242

Db        415 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 474

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Db        475 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 534

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334

Db        535 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 571


RESULT 8
Q96GA6
ID   Q96GA6      PRELIMINARY;      PRT;   614 AA.
AC   Q96GA6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Unknown (protein for MGC:15420).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-CELL;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC009851; AAH09851.1; -.
DR   InterPro; IPR000005; HTHAraC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00408; IGc2; 2.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Immunoglobulin domain.
SQ   SEQUENCE   614 AA;  67921 MW;  55EF536E77AA9BBB CRC64;

  Query Match             26.3%;  Score 488.5;  DB 4;  Length 614;
  Best Local Similarity   31.2%;  Pred. No. 4.3e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69

Db        242 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 297

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123

Db        298 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 355

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183

Db        356 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 415

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242

Db        416 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 475

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Db        476 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 535

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334

Db        536 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 572


RESULT 9
Q96AA6
ID   Q96AA6      PRELIMINARY;      PRT;   618 AA.
AC   Q96AA6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
```

```
DE   Hypothetical 67.8 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPH;
RA   Strausberg R.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC017356; AAH17356.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   618 AA;  67758 MW;  96DBD4C7C696E0A6 CRC64;

  Query Match           26.3%;  Score 488.5;  DB 4;  Length 618;
  Best Local Similarity  31.2%;  Pred. No. 4.4e-37;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
Db        246 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 301

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
Db        302 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 359

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
Db        360 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 419

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
Db        420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 479

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
Db        480 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 539

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
Db        540 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 576


RESULT 10
Q95M34
ID   Q95M34      PRELIMINARY;      PRT;   337 AA.
AC   Q95M34;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Immunogobulin gamma 1 heavy chain constant region
DE   (Fragment).
GN   IGHC1.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wagner B.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98383416; PubMed=9717671;
RA   Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA   Leibold W., Radbruch A.;
RT   "Organization of the equine immunoglobulin heavy chain constant region
RT   genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL   Immunobiology 199:105-119(1998).
DR   EMBL; AJ300675; CAC44624.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT   NON_TER       1       1
SQ   SEQUENCE   337 AA;  37438 MW;  A60BF2B01DEFD1F6 CRC64;

  Query Match           24.9%;  Score 463;  DB 6;  Length 337;
  Best Local Similarity  31.0%;  Pred. No. 4.7e-35;
  Matches  107;  Conservative   71;  Mismatches  121;  Indels   46;  Gaps   12;

Qy         23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPK 78
Db          6 PKVFALAPGCGTTSD--STVALGCLVSGYFPEPVKVSWNSGSLTSG---VHTFPSVL--- 57

Qy         79 REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSR------------------ 120
Db         58 -QSSGFYSLSSMVTVPASTW-TSETYICNVVHAASNFKVDKRIEPIPDNHQKVCDMSKCP 115

Qy        121 RCSDDE-PRGVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGS 177
Db        116 KCPAPELLGGPSVFIFPPNPKDTLMITRTPEVTCVVVDVSQENPDVKFNWYMDGVEVRTA 175

Qy        178 ASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPE 237
Db        176 TTRPKEEQFNSTYRVVSVLRIQHQDWLSGKEFKCKVNNQALPQPIERTITKTKGRSQEPQ 235

Qy        238 VYMLPPSPEETGTTR-TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPS 296
Db        236 VYVLAPHPDELSKSKVSVTCLVKDFYPPEINIEWQSNGQPELETKYSTTQAQQD--SDGS 293

Qy        297 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
Db        294 YFLYSKLSVDRNRWQQGTTFTCGVMHEALHN--------HYTQKN 330


RESULT 11
Q8VCX7
ID   Q8VCX7      PRELIMINARY;      PRT;   613 AA.
AC   Q8VCX7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 67.9 kDa protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SALIVARY GLAND;
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC018315; AAH18315.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 4.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW   Hypothetical protein.
SQ   SEQUENCE   613 AA;  67855 MW;  41A9384DD4C22862 CRC64;

  Query Match           24.5%;  Score 455.5;  DB 11;  Length 613;
  Best Local Similarity  32.1%;  Pred. No. 5.3e-34;
  Matches  105;  Conservative   63;  Mismatches  140;  Indels   19;  Gaps    8;

Qy         23 PTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLF---PYTT 75
```

```
Db        249 PNVNVF---VPPRDGFSGPAPRKSKLICEATNFTPKPITVSWLKDGKLVESGFTTDPVTI 305

Qy         76 RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCSDDEPRGVITYL 134

Db        306 ENKGSTPQTYKVISTLTISEIDWLNLNVYTCRVDHRGLTFLKNVSSTCAASPSTDILTFT 365

Qy        135 IPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITS 194

Db        366 IPPSFADIFLSKSANLTCLVSNLATYETLNISWASQSGEPLETKIKIMESHPNGTFSAKG 425

Qy        195 ILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGK--RLAPEVYMLPPSPEETG--T 250

Db        426 VASVCVEDWNNRKEFVCTVTHRDLPSPQKKFISK-PNEVHKHPPAVYLLPPAREQLNLRE 484

Qy        251 TRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIW 310

Db        485 SATVTCLVKGFSPADISVQWLQRGQLLPQEKYVTSAPMPEPGAPGFYFTHSILTVTEEEW 544

Qy        311 EKGNLVTCRVVHEALP---GSRTLEKS 334

Db        545 NSGETYTCVVGHEALPHLVTERTVDKS 571


RESULT 12
Q8TC63
ID   Q8TC63      PRELIMINARY;      PRT;   473 AA.
AC   Q8TC63;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 52.0 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025985; AAH25985.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   473 AA;  51986 MW;  E29920B09BA369F5 CRC64;

  Query Match             24.5%;  Score 455;  DB 4;  Length 473;
  Best Local Similarity   33.4%;  Pred. No. 4.2e-34;
  Matches  115;  Conservative   67;  Mismatches  128;  Indels   34;  Gaps   12;

Qy         18 VTIIPPTVK---LFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENL 70

Db        142 VSVSPASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHT 198

Qy         71 FPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEP--- 127

Db        199 FPAVL----QSSGLYSLSSVVTVPSSS-LGTKTYTCNVDHKPSNTKVDKRVESKYGPPCP 253

Qy        128 --------RGVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGS 177

Db        254 SCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVHNA 313

Qy        178 ASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPE 237

Db        314 KTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQ 373

Qy        238 VYMLPPSPEE-TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPS 296

Db        374 VYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN--NYKTTPPVLD--SDGS 429

Qy        297 FFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340

Db        430 FFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYT-QKSLSLSLG 472


RESULT 13
Q8TC77
ID   Q8TC77      PRELIMINARY;      PRT;   471 AA.
AC   Q8TC77;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 51.8 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPLEEN;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024289; AAH24289.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   471 AA;  51791 MW;  388F7F4CF588660E CRC64;

  Query Match             23.8%;  Score 441.5;  DB 4;  Length 471;
  Best Local Similarity   31.7%;  Pred. No. 7.6e-33;
  Matches  110;  Conservative   73;  Mismatches  127;  Indels   37;  Gaps   13;

Qy         18 VTIIPPTVK---LFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENL 70

Db        137 VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG---VHT 193

Qy         71 FPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE----------DSSR 120

Db        194 FPAVL----QSSGLYSLSSVVTVPSSS-LGTQTYICNVNHKPSNTKVDKKVEPKSCDKTH 248

Qy        121 RC----SDDEPRGVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKS 174

Db        249 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 308

Qy        175 IGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL 234

Db        309 HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 368

Qy        235 APEVYMLPPSPEE-TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGT 293

Db        369 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN--NYKTTPPVLD--S 424

Qy        294 DPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340

Db        425 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT-QKSLSLSPG 470


RESULT 14
Q9R1A4
ID   Q9R1A4      PRELIMINARY;      PRT;   437 AA.
AC   Q9R1A4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Gammal heavy chain of Mab7 (Fragment).
GN   IGH-4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT   "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT   antibody (Mab 7, its light and heavy chains) and construction of a
RT   single chain antibody (scFV).";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF152372; AAD40243.1; -.
DR   HSSP; P01842; 7FAB.
DR   MGD; MGI:96446; Igh-4.
```

```
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00410; IG_like; 2.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     437    437
SQ   SEQUENCE   437 AA;  48142 MW;  5C3A7BB3EE7D697C CRC64;

  Query Match             23.4%;  Score 435;  DB 11;  Length 437;
  Best Local Similarity   31.2%;  Pred. No. 2.8e-32;
  Matches  109;  Conservative   66;  Mismatches  134;  Indels   40;  Gaps   13;

Qy         16 GPVTII--------PPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVD 63
              || |::         ||:|  :  :       :| : | ||| |: |  | |||    |
Db        104 GPGTLVTVSAAKTTPPSV--YPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSS 161

Qy         64 GQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-------- 115
              |    : ||   :        ::| | | :    |  | | ||:| |  |
Db        162 G---VHTFPAVLQ-----SDLYTLSSSVTVPSSTW-PSETVTCNVAHPASSTKVDKKIVP 212

Qy        116 -EDSSRRCSDDEPRGVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERK 172
              :   : |    |     :: || | | |    |||:||:|:|:  :: |  :|  :
Db        213 RDCGCKPCICTVPEVSSVFIFPPKPKDVLTITLTPKVTCVVVDISKDDPEVQFSWFVDDV 272

Qy        173 KSIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGK 232
              :   : :|   :  ::|    | ||:  :||: |: ::|||:   || || ::|:|  |:
Db        273 EVHTAQTQPREEQFNSTFRSVSELPIMHQDWLNGKEFKCRVNSAAFPAPIEKTISKTKGR 332

Qy        233 RLAPEVYMLPPSPEETGTTR-TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDH 291
                ||:|| :||  |:    : ::||:|  |:| :|:|:|  | :     ::  |:|  |
Db        333 PKAPQVYTIPPPKEQMAKDKVSLTCMITDFFPEDITVEWQWNGQP--AENYKNTQPIMD- 389

Qy        292 GTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              || |:|:||:: | || || ||  || |:|| |    | ||:| :| |
Db        390 -TDGSYFVYSKLNVQKSNWEAGNTFTCSVLHEGLHNHHT-EKNLSHSPG 436


RESULT 15
Q99LC4
ID   Q99LC4      PRELIMINARY;      PRT;   463 AA.
AC   Q99LC4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to RIKEN cDNA 1810060O09 gene.
GN   IGH-4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC003435; AAH03435.1; -.
DR   HSSP; P01842; 7FAB.
DR   MGD; MGI:96446; Igh-4.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 2.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00410; IG_like; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ   SEQUENCE   463 AA;  51007 MW;  EAA674C6BBC30783 CRC64;

  Query Match             23.4%;  Score 434.5;  DB 11;  Length 463;
  Best Local Similarity   31.6%;  Pred. No. 3.4e-32;
  Matches  110;  Conservative   65;  Mismatches  136;  Indels   37;  Gaps   13;

Qy          9 TLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDG 64
              |:|  :: | :: |    |   |       :| : | ||| |: |  | |||    | |
Db        136 TVSAAKTTPPSVYP----LAPGSA---AQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSG 188

Qy         65 QEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF--------- 115
                 : ||   :        ::| | | :    |  | | ||:| |  |
Db        189 ---VHTFPAVLQ-----SDLYTLSSSVTVPSSTW-PSETVTCNVAHPASSTKVDKKIVPR 239

Qy        116 EDSSRRCSDDEPRGVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKK 173
              :   : |    |     :: || | | |    |||:||:|:|:  :: |  : :|  :  :
Db        240 DCGCKPCICTVPEVSSVFIFPPKPKDVLTITLTPKVTCVVVDISKDDPEVQFSWFVDDVE 299

Qy        174 SIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKR 233
              : :|   :  ::|    | ||:  :||: |: ::|||:   || || ::|:|  |:
Db        300 VHTAQTQPREEQFNSTFRSVSELPIMHQDWLNGKEFKCRVNSAAFPAPIEKTISKTKGRP 359

Qy        234 LAPEVYMLPPSPEETGTTR-TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHG 292
              ||:|| :||  |:    : ::||:|  |:| :|:|:|  | :     ::  |:|  |
Db        360 KAPQVYTIPPPKEQMAKDKVSLTCMITDFFPEDITVEWQWNGQP--AENYKNTQPIMD-- 415

Qy        293 TDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              || |:|:||:: | || || ||  || |:|| |    | ||:| :| |
Db        416 TDGSYFIYSKLNVQKSNWEAGNTFTCSVLHEGLHNHHT-EKSLSHSPG 462


Search completed: July  9, 2003, 13:41:03
Job time : 34 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 11 Seconds
(without alignments)
1285.766 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHHTLSLPESGPVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1051	56.6	429	1	EPC_RAT	P01855 rattus norv
2	940.5	50.6	421	1	EPC_MOUSE	P06336 mus musculu
3	745	40.1	428	1	EPC_HUMAN	P01854 homo sapien
4	499	26.9	454	1	MUC_HUMAN	P01871 homo sapien
5	490	26.4	479	1	MUCM_RABIT	P04221 oryctolagus
6	488	26.3	458	1	MUC_RABIT	P03988 oryctolagus
7	484.5	26.1	391	1	MUCB_HUMAN	P04220 homo sapien
8	468	25.2	326	1	GC2_HUMAN	P01859 homo sapien
9	464.5	25.0	450	1	MUC_CANFA	P01874 canis famil
10	455.5	24.5	455	1	MUC_MOUSE	P01872 mus musculu
11	455.5	24.5	476	1	MUCM_MOUSE	P01873 mus musculu
12	451.5	24.3	327	1	GC4_HUMAN	P01861 homo sapien
13	449.5	24.2	457	1	MUC_SUNMU	P20768 suncus muri
14	440.5	23.7	330	1	GC1_HUMAN	P01857 homo sapien
15	433.5	23.3	454	1	MUC_MESAU	P06337 mesocricetu
16	426	22.9	322	1	GCA_RAT	P20760 rattus norv
17	424.5	22.8	329	1	GCC_RAT	P20762 rattus norv
18	424.5	22.8	335	1	GCAB_MOUSE	P01864 mus musculu
19	423	22.8	326	1	GC1_RAT	P20759 rattus norv
20	422	22.7	323	1	GC_RABIT	P01870 oryctolagus
21	419	22.6	329	1	GC2_CAVPO	P01862 cavia porce
22	410	22.1	324	1	GC1_MOUSE	P01868 mus musculu
23	410	22.1	393	1	GC1M_MOUSE	P01869 mus musculu
24	408.5	22.0	329	1	GC3_MOUSE	P22436 mus musculu
25	405	21.8	336	1	GCB_MOUSE	P01866 mus musculu
26	405	21.8	405	1	GCBM_MOUSE	P01867 mus musculu
27	403.5	21.7	398	1	GC3M_MOUSE	P03987 mus musculu
28	394	21.2	330	1	GCAA_MOUSE	P01863 mus musculu
29	394	21.2	399	1	GCAM_MOUSE	P01865 mus musculu
30	386	20.8	290	1	GC3_HUMAN	P01860 homo sapien
31	386	20.8	438	1	HVC2_HETFR	P23085 heterodontu
32	381.5	20.5	333	1	GCB_RAT	P20761 rattus norv
33	373	20.1	438	1	HVCS_HETFR	P23087 heterodontu
34	373	20.1	461	1	HVCM_HETFR	P23088 heterodontu
35	363	19.5	393	1	HVC3_HETFR	P23086 heterodontu
36	355	19.1	340	1	ALC2_HUMAN	P01877 homo sapien
37	348	18.7	370	1	HVC1_HETFR	P23084 heterodontu
38	346.5	18.6	353	1	ALC1_HUMAN	P01876 homo sapien
39	336.5	18.1	353	1	ALC1_GORGO	P20758 gorilla gor
40	334.5	18.0	446	1	MUC_CHICK	P01875 gallus gall
41	323	17.4	344	1	ALC_MOUSE	P01878 mus musculu
42	270.5	14.6	481	1	MUCM_ICTPU	P23735 ictalurus p
43	237.5	12.8	299	1	ALC_RABIT	P01879 oryctolagus
44	210	11.3	383	1	DTC_HUMAN	P01880 homo sapien
45	160	8.6	103	1	LAC_CHICK	P20763 gallus gall

ALIGNMENTS

RESULT 1
EPC_RAT
ID EPC_RAT STANDARD; PRT; 429 AA.
AC P01855;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC STRAIN=LOU/C/WSL;
RX MEDLINE=83064537; PubMed=6292865;
RA Hellman L., Pettersson U., Engstroem A., Karlsson T., Bennich H.;
RT "Structure and evolution of the heavy chain from rat immunoglobulin
RT E.";
RL Nucleic Acids Res. 10:6041-6049(1982).
RN [2]
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6820340;
RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RT construction, identification, and DNA sequence.";
RL DNA 1:335-343(1982).
RN [3]
RP SEQUENCE OF 205-306 FROM N.A.
RX MEDLINE=82174576; PubMed=6803238;
RA Hellman L., Pettersson U., Bennich H.;
RT "Characterization and molecular cloning of the mRNA for the heavy
RT (epsilon) chain of rat immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).
CC ---

CC ---
DR EMBL; J00744; AAA41379.1; ALT_INIT.
DR PIR; A02143; EHRT.
DR HSSP; P01854; 1IGE.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00410; IG_like; 3.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT CONFLICT 168 168 R -> N (IN REF. 2).

```
FT   CONFLICT    308    308       P -> L (IN REF. 2).
SQ   SEQUENCE   429 AA;  48671 MW;  D2970B34EF8A72B0 CRC64;

  Query Match           56.6%;  Score 1051;  DB 1;  Length 429;
  Best Local Similarity   62.6%;  Pred. No. 3.3e-72;
  Matches  206;  Conservative   35;  Mismatches   78;  Indels   10;  Gaps    3;

Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              || |  ||| | ||||||    |||||| | | |      | : ||:| ::      | |
Db         97 PVNITKPTVDLLHSSCDPNA-FHSTIQLYCFVYGHIQNDVSIHWLMDDRKI-----YETH 150

Qy         77 PK----REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVIT 132
               :    :| |:  |  | :|||| |||| :|:|| |   |  :   :||||||||||||
Db        151 AQNVLIKEEGKLASTYSRLNITQQQWMSESTFTCKVTSQGENYWAHTRRCSDDEPRGVIT 210

Qy        133 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSI 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        211 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHNATTSI 270

Qy        193 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR 252
              |||||||||||||||||||||||||||||||| |||| |||||:  |  ||    |
Db        271 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKAPGKRSAPEVYVFLPPEEEEKDKR 330

Qy        253 TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEK 312
              |:||||: |:| :|||||| :::      |:|| | | :|::  ||::||: | |::| :
Db        331 TLTCLIQNFFPEDISVQWLQDSKLIPKSQHSTTTPLKYNGSNQRFFIFSRLEVTKALWTQ 390

Qy        313 GNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
                  ||||:||||   | ||:::  | ||
Db        391 TKQFTCRVIHEALREPRKLERTISKSLGN 419


RESULT 2
EPC_MOUSE
ID   EPC_MOUSE      STANDARD;      PRT;   421 AA.
AC   P06336; P01856;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig epsilon chain C region.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84236092; PubMed=6329728;
RA   Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT   "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT   comparison with the human epsilon gene sequence.";
RL   EMBO J. 1:1117-1123(1982).
RN   [2]
RP   REVISIONS.
RA   Honjo T.;
RL   Submitted (APR-1986) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 34-421 FROM N.A.
RX   MEDLINE=83117774; PubMed=6818553;
RA   Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT   "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
RT   chain cDNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01857; CAA25977.1; -.
DR   EMBL; X01857; CAA25978.1; -.
DR   PIR; A02145; EHMSS.
DR   PIR; A02144; EHMS.
DR   HSSP; P01854; 1IGE.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN        1     90       CH1.
FT   DOMAIN       91    197       CH2.
FT   DOMAIN      198    304       CH3.
FT   DOMAIN      305    421       CH4.
FT   DISULFID     23     75       BY SIMILARITY.
FT   DISULFID    121    180       BY SIMILARITY.
FT   DISULFID    226    285       BY SIMILARITY.
FT   DISULFID    330    392       BY SIMILARITY.
FT   CARBOHYD     43     43       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     72     72       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     84     84       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     95     95       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    238    238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    261    261       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    365    365       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    415    415       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   421 AA;  47320 MW;  8F909E1F30A06B47 CRC64;

  Query Match           50.6%;  Score 940.5;  DB 1;  Length 421;
  Best Local Similarity   55.2%;  Pred. No. 6.9e-64;
  Matches  191;  Conservative   46;  Mismatches  100;  Indels    9;  Gaps    5;

Qy          2 FHHHHHHTLSLPES-----GPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKV 56
              |  |  |  |  ||      || |  ||::| ||||||    ||||||  | : |     |
Db         73 FTCHVTHPPSFNESRTILVRPVNITEPTLELLHSSCDPNA-FHSTIQLYCFIYGHILNDV 131

Qy         57 HVTWLVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
               |:||:| :|  :     |   | | |:  |  |::||| ||||  :|:|| |   |  :
Db        132 SVSWLMDDREITDTLAQTVLIKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYL 190

Qy        117 DSSRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIG 176
              :||| | |||||||||||||||:|| |||||||:|||||:|: ||| :|:| |:
Db        191 AHTRRCPDHEPRGVITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV- 249

Qy        177 SASQRSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLA 235
              ||||  |||| :|||||||||| |||||| |||| |||| |||||||||| ||:| |
Db        250 SASQWYTKHHNNATTSITSILPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSA 309

Qy        236 PEVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDP 295
              ||||: ||  ||:   ||:|||: |:| :|||||| : :      |:|| | | :|::
Db        310 PEVYVFPPPEEESEDKRTLTCLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNQ 369

Qy        296 SFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              ||::||: | |::| :    ||:|:||||   | |||::  | ||
Db        370 GFFIFSRLEVAKTLWTQRKQFTCQVIHEALQKPRKLEKTISTSLGN 415


RESULT 3
EPC_HUMAN
ID   EPC_HUMAN      STANDARD;      PRT;   428 AA.
AC   P01854;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ig epsilon chain C region.
GN   IGHE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83168897; PubMed=6300763;
RA   Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,
RA   Kikuchi M., Sugino Y., Nishida Y., Honjo T.;
RT   "Molecular cloning and nucleotide sequencing of human immunoglobulin
RT   epsilon chain cDNA.";
RL   Nucleic Acids Res. 11:719-726(1983).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT LEU-359.
RX   MEDLINE=83001945; PubMed=6288268;
RA   Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;
RT   "Duplication and deletion in the human immunoglobulin epsilon genes.";
RL   Cell 29:691-699(1982).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84236029; PubMed=6234164;
RA   Flanagan J.G., Rabbitts T.H.;
RT   "The sequence of a human immunoglobulin epsilon heavy chain constant
RT   region gene, and evidence for three non-allelic genes.";
RL   EMBO J. 1:655-660(1982).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84207910; PubMed=6327276;
RA   Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;
RT   "Long terminal repeat-like elements flank a human immunoglobulin
RT   epsilon pseudogene that lacks introns.";
RL   EMBO J. 1:1539-1544(1982).
RN   [5]
RP   PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).
RA   Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
RL   (In) Bach M.K. (eds.);
RL   Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL   Marcel Dekker, New York (1978).
RN   [6]
RP   SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.
RX   MEDLINE=83065234; PubMed=6815656;
RA   Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,
RA   Bell L.O., Gould H.J.;
RT   "Cloning and sequence determination of the gene for the human
RT   immunoglobulin epsilon chain expressed in a myeloma cell line.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
RN   [7]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=87089848; PubMed=3796618;
RA   Padlan E.A., Davies D.R.;
RT   "A model of the Fc of immunoglobulin E.";
RL   Mol. Immunol. 23:1063-1075(1986).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L00022; AAB59424.1; ALT_INIT.
DR   PIR; A02142; EHHU.
DR   PIR; A22771; A22771.
DR   PIR; A23195; A23195.
DR   PDB; 1IGE; 15-JUL-92.
DR   Genew; HGNC:5522; IGHE.
DR   MIM; 147180; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00407; IGc1; 4.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW   3D-structure.
FT   NON_TER       1      1
FT   DISULFID     14     14       INTERCHAIN (WITH A LIGHT CHAIN).
FT   DISULFID     15    105
FT   DISULFID     29     85
FT   DISULFID    121    121       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    135    193
FT   DISULFID    209    209       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    239    299
FT   DISULFID    345    405
FT   CARBOHYD     21     21       N-LINKED (GLCNAC. . .).
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .).
FT   CARBOHYD     99     99       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    146    146       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    252    252       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    275    275       N-LINKED (GLCNAC. . .).
FT   VARIANT     359    359       W -> L (POSSIBLE POLYMORPHISM).
FT                                /FTId=VAR_003885.
SQ   SEQUENCE   428 AA;  47019 MW;  25C4CA072AA558A0 CRC64;

  Query Match             40.1%;  Score 745;  DB 1;  Length 428;
  Best Local Similarity   44.4%;  Pred. No. 3.9e-49;
  Matches  143;  Conservative   62;  Mismatches  109;  Indels    8;  Gaps    5;

Qy         22 PPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRPKREG 81
              |||||:  ||||  |    ||||||||||::|  :::||| |||  :     :|    :
Db        111 PPTVKILQSSCDGGGHFPPTIQLLCLVSGYTPGTINITWLEDGQVMD--VDLSTASTTQE 168

Qy         82 GQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIPPSPLD 141
              |:  | |||: ::|  |:|  |||| | : |  |||:::|:|  ||||  ||  ||| |
Db        169 GELASTQSELTLSQKHWLSDRTYTCQVTYQGHTFEDSTKKCADSNPRGVSAYLSRPSPFD 228

Qy        142 LYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSASQRSTKHHHATTSITSILPVDA 200
              |:   :| :||||:||  |:  : :|| |   | :  ::::  |  : | ::|| |||
Db        229 LFIRKSPTITCLVVDLAPSKGTVNLTWSRASGKPVNHSTRKEEKQRNGTLTVTSTLPVGT 288

Qy        201 KDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGT--TRTVTCLI 258
              :|||||| ||||| ||| |: ::|| ||  | | |||||    :||  |:   ||: |||
Db        289 RDWIEGETYQCRVTHPHLPRALMRSTTKTSGPRAAPEVYAF-ATPEWPGSRDKRTLACLI 347

Qy        259 RGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVTC 318
              : | | :|||||| |  :     |:||:|:|  |:   ||::||: | :: ||: :   |
Db        348 QNFMPEDISVQWLHNEVQLPDARHSTTQPRKTKGS--GFFVFSRLEVTRAEWEQKDEFIC 405

Qy        319 RVVHEALPGSRTLEKSLHYSAG 340
              | ||||   |:|:::::  : |
Db        406 RAVHEAASPSQTVQRAVSVNPG 427


RESULT 4
MUC_HUMAN
ID   MUC_HUMAN      STANDARD;      PRT;   454 AA.
AC   P01871;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ig mu chain C region.
GN   IGHM.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 1-434 FROM N.A.
RX   MEDLINE=90332450; PubMed=2115996;
RA   Friedlander R.M., Nussenzweig M.C., Leder P.;
RT   "Complete nucleotide sequence of the membrane form of the human IgM
RT   heavy chain.";
RL   Nucleic Acids Res. 18:4278-4278(1990).
RN   [2]
RP   SEQUENCE (WALDENSTROM'S MACROGLOBULIN GAL).
RX   MEDLINE=75059123; PubMed=4803843;
RA   Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
RT   "The primary structure of a monoclonal IgM-immunoglobulin
RT   (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
RT   type), subgroup H III. Architecture of the complete IgM-molecule.";
```

```
RL   Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
RN   [3]
RP   REVISIONS (GAL).
RX   MEDLINE=81066716; PubMed=6777162;
RA   Mihaesco E., Barnikol-Watanabe S., Barnikol H.U., Mihaesco C.,
RA   Hilschmann N.;
RT   "The primary structure of the constant part of mu-chain-disease
RT   protein BOT.";
RL   Eur. J. Biochem. 111:275-286(1980).
RN   [4]
RP   SEQUENCE (WALDENSTROM'S OU), DISULFIDE BONDS, AND CARBOHYDRATES.
RX   MEDLINE=74005511; PubMed=4742735;
RA   Putnam F.W., Florent G., Paul C., Shinoda T., Shimizu A.;
RT   "Complete amino acid sequence of the Mu heavy chain of a human IgM
RT   immunoglobulin.";
RL   Science 182:287-291(1973).
RN   [5]
RP   PARTIAL SEQUENCE FROM N.A.
RX   MEDLINE=82059479; PubMed=6795593;
RA   Rabbitts T.H., Forster A., Milstein C.P.;
RT   "Human immunoglobulin heavy chain genes: evolutionary comparisons of
RT   C mu, C delta and C gamma genes and associated switch sequences.";
RL   Nucleic Acids Res. 9:4509-4524(1981).
RN   [6]
RP   SEQUENCE OF 299-387 AND 438-454 FROM N.A.
RX   MEDLINE=81077306; PubMed=6777778;
RA   Dolby T.W., Devuono J., Croce C.M.;
RT   "Cloning and partial nucleotide sequence of human immunoglobulin mu
RT   chain cDNA from B cells and mouse-human hybridomas.";
RL   Proc. Natl. Acad. Sci. U.S.A. 77:6027-6031(1980).
CC   -!- MISCELLANEOUS: ALL 4 COMBINATIONS OF THE S/G & V/G POLYMORPHISMS
CC       AT POSITIONS 192 AND 216 HAVE BEEN OBSERVED IN HUMAN MU CHAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17115; CAA34971.1; ALT_SEQ.
DR   EMBL; X57086; -; NOT_ANNOTATED_CDS.
DR   PIR; A02162; MHHU.
DR   HSSP; P01857; 1FC1.
DR   Genew; HGNC:5541; IGHM.
DR   MIM; 147020; -.
DR   GlycoSuiteDB; P01871; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00407; IGC1; 3.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN        1    105       CH1.
FT   DOMAIN      106    218       CH2.
FT   DOMAIN      219    324       CH3.
FT   DOMAIN      325    454       CH4.
FT   DISULFID     14     14       INTERCHAIN (WITH A LIGHT CHAIN).
FT   DISULFID     28     88
FT   DISULFID    135    198
FT   DISULFID    215    215       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    245    304
FT   DISULFID    292    292       INTERCHAIN (WITH A HEAVY CHAIN IN ANOTHER
FT                                OF THE 5 TETRAMERIC SUBUNITS OF THE
FT                                MOLECULE).
FT   DISULFID    352    414
FT   DISULFID    453    453       INTERCHAIN (WITH A HEAVY CHAIN).
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    210    210       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    273    273       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    441    441       N-LINKED (GLCNAC. . .).
FT                                /FTId=CAR_000219.
FT   VARIANT     192    192       S -> G.
FT                                /FTId=VAR_003903.
FT   VARIANT     216    216       V -> G.
FT                                /FTId=VAR_003904.
SQ   SEQUENCE   454 AA;  49556 MW;  21EC72EADC56922E CRC64;

  Query Match             26.9%;  Score 499;  DB 1;  Length 454;
  Best Local Similarity   31.5%;  Pred. No: 1.5e-30;
  Matches  106;  Conservative   72;  Mismatches  135;  Indels   24;  Gaps    9;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|   || | :|      ||    |:  |  :|:|  :|||| :: .|:|| :|:: :
Db        102 PLPVIAELPPKVSVF---VPPRDGFFGNPRSKSKLICQATGFSPRQIQVSWLREGKQVGS 158

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||          |  |  |: : | : | :  |:| : :|| | | |  | :::|  |
Db        159 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 216

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
               |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        217 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 276

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              .  | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        277 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 336

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:  :      : |: |  :      :| :
Db        337 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 396

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        397 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 433


RESULT 5
MUCM_RABIT
ID   MUCM_RABIT     STANDARD;      PRT;   479 AA.
AC   P04221;
DT   20-MAR-1987 (Rel. 04, Created)  .
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig mu chain C region membrane-bound form.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A. (A2 ALLOTYPE).
RX   MEDLINE=84088930; PubMed=6418803;
RA   Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;
RT   "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain
RT   of VHa2 allotype: comparisons with VHa1 and membrane mu sequences.";
RL   J. Immunol. 132:490-495(1984).
CC   -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC       SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC       THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC       TERMINAL SEGMENTS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K01357; AAA31293.1; -.
DR   PIR; A02165; MHRBM.
```

```
DR   HSSP; P01842; 7FAB.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW   Alternative splicing; Transmembrane.
FT   NON_TER       1      1
FT   DOMAIN        1    106       CH1.
FT   DOMAIN      107    222       CH2.
FT   DOMAIN      223    327       CH3.
FT   DOMAIN      328    458       CH4.
FT   TRANSMEM    459    476       POTENTIAL.
FT   DISULFID     14     14       INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT   DISULFID     28     90       BY SIMILARITY.
FT   DISULFID    137    200       BY SIMILARITY.
FT   DISULFID    219    219       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    249    308       BY SIMILARITY.
FT   DISULFID    296    296       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    356    418       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    212    212       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    261    261       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   479 AA;  52351 MW;  689C637A47BE19FC CRC64;

  Query Match             26.4%;  Score 490;  DB 1;  Length 479;
  Best Local Similarity   32.4%;  Pred. No. 7.7e-30;
  Matches  115;  Conservative   62;  Mismatches  154;  Indels   24;  Gaps    9;

Qy          5 HHHHTLSLPESGPV-TIIPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVT 59
              | :    |  | || : :|| | :|     ||    |      :|:|  :|||| :: |:
Db         94 HSNSNRDLRVSFPVDSELPPNVSVF---IPPRDSFSGSGTRKSRLICQATGFSPKQISVS 150

Qy         60 WLVDGQEAEN---LFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
              || |||: |:     |     |  |  |||: | : ||:  |:| : ||| | | |  |:
Db        151 WLRDGQKVESGVLTKPVEAETKGAGPATFSISSMLTITESDWLSQSLYTCRVDHRGIFFD 210

Qy        117 DS---SRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKK 173
               :   |  ||      |:  : | ||  | : : : :| ||| || :  :: ::|     |
Db        211 KNVSMSSECSTTPSPGIQVFPIAPSFADTFLSKSARLICLVTDLTTYGSLNISWASHNGK 270

Qy        174 SIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKR 233
              :: :    :  | :|| |      | |:||  || : | | |   | |:  :|:|   :
Db        271 ALDTHMNITESHPNATFSAMGEASVCAEDWESGEQFTCTVTHADLPFPLKHTISK--SRE 328

Qy        234 LA---PEVYMLPPSPEE--TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQ 288
              :|   | ||:|||: |:     : |||||::|| |::: |||    :   :  : |: |
Db        329 VAKHPPAVYVLPPAREQLVLRESATVTCLVKGFSPADVFVQWQQRGQPLSSDKYVTSAPA 388

Qy        289 KDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALP---GSRTLEKSLHYSAG 340
              :      :| :| : | :  |  |   || | |||||     ||::||      |
Db        389 PEPQAPGLYFTHSTLTVTEEDWNSGETFTCVVGHEALPHMVTERTVDKSTEGEVG 443


RESULT 6
MUC_RABIT
ID   MUC_RABIT      STANDARD;      PRT;   458 AA.
AC   P03988;
DT   23-OCT-1986 (Rel. 02, Created)
DT   23-OCT-1986 (Rel. 02, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig mu chain C region.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A. (A2 ALLOTYPE).
RX   MEDLINE=84088930; PubMed=6418803;
RA   Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;
RT   "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain
RT   of VHa2 allotype: comparisons with VHa1 and membrane mu sequences.";
RL   J. Immunol. 132:490-495(1984).
CC   -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC       SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC       THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC       TERMINAL SEGMENTS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; K01357; -; NOT_ANNOTATED_CDS.
DR   PIR; A02164; MHRB.
DR   HSSP; P01842; 7FAB.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW   Alternative splicing.
FT   NON_TER       1      1
FT   DOMAIN        1    106       CH1.
FT   DOMAIN      107    222       CH2.
FT   DOMAIN      223    327       CH3.
FT   DOMAIN      328    458       CH4.
FT   DISULFID     14     14       INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT   DISULFID     28     90       BY SIMILARITY.
FT   DISULFID    137    200       BY SIMILARITY.
FT   DISULFID    219    219       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    249    308       BY SIMILARITY.
FT   DISULFID    296    296       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    356    418       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    212    212       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    261    261       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   458 AA;  49897 MW;  F338D6A3B28E4864 CRC64;

  Query Match             26.3%;  Score 488;  DB 1;  Length 458;
  Best Local Similarity   32.7%;  Pred. No. 1e-29;
  Matches  114;  Conservative   62;  Mismatches  149;  Indels   24;  Gaps    9;

Qy          5 HHHHTLSLPESGPV-TIIPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVT 59
              | :    |  | || : :|| | :|     ||    |      :|:|  :|||| :: |:
Db         94 HSNSNRDLRVSFPVDSELPPNVSVF---IPPRDSFSGSGTRKSRLICQATGFSPKQISVS 150

Qy         60 WLVDGQEAEN---LFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
              || |||: |:     |     |  |  |||: | : ||:  |:| : ||| | | |  |:
Db        151 WLRDGQKVESGVLTKPVEAETKGAGPATFSISSMLTITESDWLSQSLYTCRVDHRGIFFD 210

Qy        117 DS---SRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKK 173
               :   |  ||      |:  : | ||  | : : : :| ||| || :  :: ::|     |
Db        211 KNVSMSSECSTTPSPGIQVFPIAPSFADTFLSKSARLICLVTDLTTYGSLNISWASHNGK 270

Qy        174 SIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKR 233
              :: :    :  | :|| |      | |:||  || : | | |   | |:  :|:|   :
Db        271 ALDTHMNITESHPNATFSAMGEASVCAEDWESGEQFTCTVTHADLPFPLKHTISK--SRE 328

Qy        234 LA---PEVYMLPPSPEE--TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQ 288
```

```
              :|   | ||:|||: |:     : |||||::|| |::: |||    :   :  : |: |
Db        329 VAKHPPAVYVLPPAREQLVLRESATVTCLVKGFSPADVFVQWQQRGQPLSSDKYVTSAPA 388

Qy        289 KDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALP---GSRTLEKS 334
              :      :| :| : | :  |  |   || | |||||     ||::||
Db        389 PEPQAPGLYFTHSTLTVTEEDWNSGETFTCVVGHEALPHMVTERTVDKS 437


RESULT 7
MUCB_HUMAN
ID   MUCB_HUMAN     STANDARD;      PRT;   391 AA.
AC   P04220;
DT   20-MAR-1987 (Rel. 04, Created)
DT   20-MAR-1987 (Rel. 04, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ig mu heavy chain disease protein (BOT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84184186; PubMed=6425189;
RA   Barnikol-Watanabe S., Mihaesco E., Mihaesco C., Barnikol H.U.,
RA   Hilschmann N.;
RT   "The primary structure of mu-chain-disease protein BOT. Peculiar
RT   amino-acid sequence of the N-terminal 42 positions.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 365:105-118(1984).
CC   -!- MISCELLANEOUS: THIS PROTEIN HAS NO V REGION HOMOLOGY OR CH1
CC       REGION.
DR   PIR; A02163; MHHUBT.
DR   HSSP; P01857; 1FC1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region.
FT   DOMAIN        1     42       PRE-C-PART (NO V REGION HOMOLOGY).
FT   DOMAIN       43    155       CH2.
FT   DOMAIN      156    261       CH3.
FT   DOMAIN      262    391       CH4.
SQ   SEQUENCE   391 AA;  43057 MW;  9100843AF0CF021A CRC64;

  Query Match             26.1%;  Score 484.5;  DB 1;  Length 391;
  Best Local Similarity   31.2%;  Pred. No. 1.6e-29;
  Matches  105;  Conservative   72;  Mismatches  137;  Indels   23;  Gaps   10;

Qy         15 SGPVTI-IPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              | ||   :|| | :|     ||    |:   : :|:|  :|||| :: |:|| :|::  :
Db         40 SQPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIEVSWLREGKQVGS 95

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||       |  |  |: : | : | :  |:  : :|| | | |  | :::|  |
Db         96 --GVTTDEVEAEAKESGPTTYKVTSTLTIKESDWLGQSMFTCRVDHRGLTFQQNASSMCG 153

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
               |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        154 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQDGEAVKTHTNISE 213

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
               | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        214 SHPNATFSAVGEASICEDDWDSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 273

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
               |: |:     : |:|||: || |::: |||:   :      : |: |  :      :| :
Db        274 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 333

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
               | : |::  |  |   || | |||||     ||::||
Db        334 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 370


RESULT 8
GC2_HUMAN
ID   GC2_HUMAN      STANDARD;      PRT;   326 AA.
AC   P01859;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ig gamma-2 chain C region.
GN   IGHG2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE OF 2-326 FROM N.A.
RX   MEDLINE=82197621; PubMed=6804948;
RA   Ellison J.W., Hood L.E.;
RT   "Linkage and sequence homology of two human immunoglobulin gamma
RT   heavy chain constant region genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
RN   [2]
RP   SEQUENCE OF 88-115 FROM N.A.
RC   TISSUE=Fetal liver;
RX   MEDLINE=83001943; PubMed=6811139;
RA   Takahashi N., Ueda S., Obata M., Nikaido T., Nakai S., Honjo T.;
RT   "Structure of human immunoglobulin gamma genes: implications for
RT   evolution of a gene family.";
RL   Cell 29:671-679(1982).
RN   [3]
RP   SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RC   TISSUE=Fetal liver;
RX   MEDLINE=84235992; PubMed=6329676;
RA   Krawinkel U., Rabbitts T.H.;
RT   "Comparison of the hinge-coding segments in human immunoglobulin gamma
RT   heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
RT   genes.";
RL   EMBO J. 1:403-407(1982).
RN   [4]
RP   SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX   MEDLINE=81007873; PubMed=6774012;
RA   Wang A.-C., Tung E., Fudenberg H.H.;
RT   "The primary structure of a human IgG2 heavy chain: genetic,
RT   evolutionary, and functional implications.";
RL   J. Immunol. 125:1048-1054(1980).
RN   [5]
RP   SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
RX   MEDLINE=80001357; PubMed=113060;
RA   Connell G.E., Parr D.M., Hofmann T.;
RT   "The amino acid sequences of the three heavy chain constant region
RT   domains of a human IgG2 myeloma protein.";
RL   Can. J. Biochem. 57:758-767(1979).
RN   [6]
RP   SEQUENCE OF 238-275 (ZIE).
RX   MEDLINE=80114419; PubMed=118920;
RA   Hofmann T., Parr D.M.;
RT   "A note of the amino acid sequence of residues 381-391 of human
RT   immunoglobulins gamma chains.";
RL   Mol. Immunol. 16:923-925(1979).
RN   [7]
RP   REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
RA   Hofmann T., Parr D.M.;
RL   Submitted (MAR-1980) to the PIR data bank.
RN   [8]
RP   SEQUENCE OF 1-121 (DOT).
RX   MEDLINE=95255298; PubMed=7737190;
RA   Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT   "Characterization of the two unique human anti-flavin monoclonal
RT   immunoglobulins.";
RL   Eur. J. Biochem. 228:886-893(1995).
RN   [9]
RP   DISULFIDE BONDS.
```

```
RX   MEDLINE=72033500; PubMed=4940472;
RA   Milstein C., Frangione B.;
RT   "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
RL   Biochem. J. 121:217-225(1971).
RN   [10]
RP   DISULFIDE BONDS.
RX   MEDLINE=69064124; PubMed=5782707;
RA   Frangione B., Milstein C., Pink J.R.L.;
RT   "Structural studies of immunoglobulin G.";
RL   Nature 221:145-148(1969).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J00230; AAB59393.1; -.
DR   PIR; A02148; G2HU.
DR   HSSP; P01857; 1FC1.
DR   Genew; HGNC:5526; IGHG2.
DR   MIM; 147110; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   Immunoglobulin domain; Immunoglobulin C region.
FT   NON_TER       1      1
FT   DOMAIN        1     98       CH1.
FT   DOMAIN       99    110       HINGE.
FT   DOMAIN      111    219       CH2.
FT   DOMAIN      220    326       CH3.
FT   DISULFID     14     14       INTERCHAIN (WITH A LIGHT CHAIN).
FT   DISULFID     27     83
FT   DISULFID    102    102       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    103    103       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    106    106       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    109    109       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    140    200
FT   DISULFID    246    304
FT   SITE        156    156       AT OR NEAR THE COMPLEMENT-BINDING SITE.
FT   MOD_RES     326    326       REMOVED POST-TRANSLATIONALLY (PROBABLE).
FT   VARIANT      60     60       S -> A (IN MYELOMA PROTEINS TIL & ZIE).
FT                                /FTId=VAR_003889.
FT   CONFLICT    109    109       C -> S (IN REF. 3).
SQ   SEQUENCE   326 AA;  35884 MW;  8310878C6878CF9C CRC64;

  Query Match             25.2%;  Score 468;  DB 1;  Length 326;
  Best Local Similarity   34.4%;  Pred. No. 2.2e-28;
  Matches  111;  Conservative   66;  Mismatches  116;  Indels   30;  Gaps   12;

Qy         35 RGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSE 90
              |  : ||  | |||  : |  | |:|     |  |    :  ||       :  :|| |
Db         16 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSV 68

Qy         91 VNITQGQWMSSNTYTCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPL 140
              |  :    :  : ||||:| |    |  : :   |:|  : |         |   :| || |
Db         69 VTVPSSNF-GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPK 127

Qy        141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPV 198
              | |  : ||:|||:|:|:  |:  :   |  :  :  . : ::   :  ::|  : |:| |
Db        128 DTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTV 187

Qy        199 DAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCL 257
                :||: |: |:|:|  :    | || ::|:|  |:   |:|| |||| || |    ::|||
Db        188 VHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCL 247

Qy        258 IRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVT 317
              ::|||||:|:|:|  | : ::  :: || |  |  :| ||||||:: |:|| |::||: :
Db        248 VKGFYPSDIAVEWESNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFS 303

Qy        318 CRVVHEALPGSRTLEKSLHYSAG 340
              | |:||||    | :|||  | |
Db        304 CSVMHEALHNHYT-QKSLSLSPG 325


RESULT 9
MUC_CANFA
ID   MUC_CANFA      STANDARD;      PRT;   450 AA.
AC   P01874;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig mu chain C region.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE OF 1-177 (MOO).
RX   MEDLINE=80077682; PubMed=117299;
RA   McCumber L.J., Capra J.D.;
RT   "The complete amino-acid sequence of a canine mu chain.";
RL   Mol. Immunol. 16:565-570(1979).
RN   [2]
RP   SEQUENCE OF 178-450 (MOO).
RX   MEDLINE=78180587; PubMed=653360;
RA   Wasserman R.L., Capra J.D.;
RT   "Amino acid sequence of the Fc region of a canine immunoglobulin M:
RT   interspecies homology for the IgM class.";
RL   Science 200:1159-1161(1978).
DR   PIR; A02169; MHDG.
DR   HSSP; P01857; 1FC1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region.
FT   NON_TER       1      1
SQ   SEQUENCE   450 AA;  48895 MW;  9D460DA9D1012F5D CRC64;

  Query Match             25.0%;  Score 464.5;  DB 1;  Length 450;
  Best Local Similarity   31.8%;  Pred. No. 6e-28;
  Matches  109;  Conservative   69;  Mismatches  136;  Indels   29;  Gaps   12;

Qy         10 LSLPE--SGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW-LVDGQE 66
              |:||   ||    |||    |   :||      ||:|  |||||  :|   | | ||::
Db        106 LTLPPEVSG---FIPPRDAFFG---BPRKS-----QLICQASGFSPRQV---WSLRDGKQ 151

Qy         67 AENLFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSR 120
              |:    ||       |  |  |: : | : | :  |:| : :|| |:|  |  | :::|
Db        152 IES--GVTTNEVZAZAKZSGPTTYKVTSMLTIQEDAWLSQSVFTCKVEHRGLTFQQNASS 209

Qy        121 RCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQ 180
              |: |:| |:  : ||||   ::   : ||:||| || : :::|::| ||   :: : :
Db        210 MCTSDQPVGISIFTIPPSFASIFNTKSAKLSCLVTDLATYDSVTISWTREENGALKTHTN 269

Qy        181 RSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVY 239
              |  | : | |      |  ::|  || : | | |    |  : ::|::  |  :  | ||
Db        270 ISESHPNGTFSAMGEATVCVEEWESGEQFTCTVTHTDLPSVLKQTISRPKGVAVHMPSVY 329

Qy        240 MLPPSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSF 297
              :|||| |:     : |::||: |: | :: |||:   :       : |: |  :      :
Db        330 VLPPSREQLDLRESATLSCLVTGYSPPDVFVQWVQKGQPVPPDSYVTSAPMPEPQAPGLY 389

Qy        298 FLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              | :| : |::  |  |   || | ||:|| :|  |:|:  | |
Db        390 FAHSILTVSEEEWNAGETYTCVVAHESLP-NRVTERSVDKSTG 431
```

```
RESULT 10
MUC_MOUSE
ID   MUC_MOUSE      STANDARD;      PRT;   455 AA.
AC   P01872;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ig mu chain C region.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81076590; PubMed=6255422;
RA   Kawakami T., Takahashi N., Honjo T.;
RT   "Complete nucleotide sequence of mouse immunoglobulin mu gene and
RT   comparison with other immunoglobulin heavy chain genes.";
RL   Nucleic Acids Res. 8:3933-3945(1980).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82051295; PubMed=6795090;
RA   Goldberg G.I., Vanin E.F., Zrolka A.M., Blattner F.R.;
RT   "Sequence of the gene for the constant region of the mu chain of
RT   Balb/c mouse immunoglobulin.";
RL   Gene 15:33-42(1981).
RN   [3]
RP   SEQUENCE FROM N.A. (MYELOMA TEPC183).
RX   MEDLINE=81165562; PubMed=6260591;
RA   Auffray C., Rougeon F.;
RT   "Nucleotide sequence of a cloned cDNA corresponding to secreted mu
RT   chain of mouse immunoglobulin.";
RL   Gene 12:77-86(1980).
RN   [4]
RP   SEQUENCE (MYELOMA PROTEIN MOPC 104E).
RX   MEDLINE=79223904; PubMed=111247;
RA   Kehry M.R., Sibley C.H., Fuhrman J.S., Schilling J.W., Hood L.E.;
RT   "Amino acid sequence of a mouse immunoglobulin mu chain.";
RL   Proc. Natl. Acad. Sci. U.S.A. 76:2932-2936(1979).
RN   [5]
RP   REVISION (MOPC 104E).
RX   MEDLINE=83075344; PubMed=6816276;
RA   Kehry M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RA   Hood L.E.;
RT   "Complete amino acid sequence of a mouse mu chain: homology among
RT   heavy chain constant region domains.";
RL   Biochemistry 21:5415-5424(1982).
CC   -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC       SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC       THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC       TERMINAL SEGMENTS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00827; -; NOT_ANNOTATED_CDS.
DR   PIR; A02166; MHMS.
DR   HSSP; P01857; 1FC1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW   Alternative splicing.
FT   NON_TER       1      1
FT   DOMAIN        1    105       CH1.
FT   DOMAIN      106    217       CH2.
FT   DOMAIN      218    324       CH3.
FT   DOMAIN      325    436       CH4.
FT   DOMAIN      437    455       C-TERMINAL SEGMENT OF SECRETED FORM.
FT   DISULFID     14     14       INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT   DISULFID     28     89
FT   DISULFID    136    199       BY SIMILARITY.
FT   DISULFID    216    216       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    246    305       BY SIMILARITY.
FT   DISULFID    293    293       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    353    415       BY SIMILARITY.
FT   DISULFID    454    454       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    243    243       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    281    281       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    442    442       N-LINKED (GLCNAC. . .).
FT   VARIANT      78     78       S -> N (IN MOPC 104E).
FT   VARIANT     101    101       H -> Q (IN MOPC 104E).
FT   VARIANT     226    226       T -> N (IN TEPC183 AND MOPC 104E).
FT   VARIANT     258    258       N -> T (IN MOPC 104E).
FT   VARIANT     258    258       N -> S (IN TEPC183).
FT   VARIANT     368    368       L -> K (IN TEPC183 AND MOPC 104E).
SQ   SEQUENCE   455 AA;  50101 MW;  4CBE57CB602F9B51 CRC64;

  Query Match             24.5%;  Score 455.5;  DB 1;  Length 455;
  Best Local Similarity   32.1%;  Pred. No. 2.9e-27;
  Matches  105;  Conservative   63;  Mismatches  140;  Indels   19;  Gaps    8;

Qy         23 PTVKLFPHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLF---PYTT 75
              | | :|     ||    | |    :|:|  : |:|  : |:|| ||:  |: |   | |
Db        112 PNVNVF---VPPRDGFSGPAPRKSKLICEATNFTPKPITVSWLKDGKLVESGFTTDPVTI 168

Qy         76 RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCSDDEPRGVITYL 134
                 |    ||: : | : |::  |:: | ||| | | |  | :: |  |:      ::|:
Db        169 ENKGSTPQTYKVISTLTISEIDWLNLNVYTCRVDHRGLTFLKNVSSTCAASPSTDILTFT 228

Qy        135 IPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITS 194
              ||||  |:: : :  ||||| :| : | : ::|  :  : : :  :    | : | |
Db        229 IPPSFADIFLSKSANLTCLVSNLATYETLNISWASQSGEPLETKIKIMESHPNGTFSAKG 288

Qy        195 ILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGK--RLAPEVYMLPPSPEETG--T 250
              :  |  :||   : : | | |   | |  : |:| | :  :  | ||:|||: |:
Db        289 VASVCVEDWNNRKEFVCTVTHRDLPSPQKKFISK-PNEVHKHPPAVYLLPPAREQLNLRE 347

Qy        251 TRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIW 310
              : |||||::|| |::||||||   :      : |: |  : |    :| :| : | :  |
Db        348 SATVTCLVKGFSPADISVQWLQRGQLLPQEKYVTSAPMPEPGAPGFYFTHSILTVTEEEW 407

Qy        311 EKGNLVTCRVVHEALP---GSRTLEKS 334
              |   || | |||||     ||::||
Db        408 NSGETYTCVVGHEALPHLVTERTVDKS 434


RESULT 11
MUCM_MOUSE
ID   MUCM_MOUSE     STANDARD;      PRT;   476 AA.
AC   P01873;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig mu chain C region membrane-bound form.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE OF 433-476 FROM N.A.
RX   MEDLINE=80222874; PubMed=6771020;
RA   Early P., Rogers J., Davis M., Calame K., Bond M., Wall R.,
```

```
RA   Hood L.;
RT   "Two mRNAs can be produced from a single immunoglobulin mu gene by
RT   alternative RNA processing pathways.";
RL   Cell 20:313-319(1980).
RN   [2]
RP   SEQUENCE OF 410-476 FROM N.A. (MYELOMA MOPC 104E).
RX   MEDLINE=80222873; PubMed=6771019;
RA   Rogers J., Early P., Carter C., Calame K., Bond M., Hood L.,
RA   Wall R.;
RT   "Two mRNAs with different 3' ends encode membrane-bound and secreted
RT   forms of immunoglobulin mu chain.";
RL   Cell 20:303-312(1980).
CC   -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC       SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC       THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC       TERMINAL SEGMENTS.
CC   -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-409 IS ASSUMED TO BE
CC       IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00821; CAA24202.1; -.
DR   PIR; A02167; MHMSM.
DR   HSSP; P01857; 1FC1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGC1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW   Alternative splicing; Transmembrane.
FT   NON_TER       1      1
FT   DOMAIN        1    105       CH1.
FT   DOMAIN      106    217       CH2.
FT   DOMAIN      218    324       CH3.
FT   DOMAIN      325    436       CH4.
FT   TRANSMEM    456    473       POTENTIAL.
FT   DISULFID     14     14       INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT   DISULFID     28     89
FT   DISULFID    136    199       BY SIMILARITY.
FT   DISULFID    216    216       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    246    305       BY SIMILARITY.
FT   DISULFID    293    293       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    353    415       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    243    243       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    281    281       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   476 AA;  52656 MW;  8D476575A5204071 CRC64;

  Query Match           24.5%;  Score 455.5;  DB 1;  Length 476;
  Best Local Similarity  32.1%;  Pred. No. 3.1e-27;
  Matches  105;  Conservative   63;  Mismatches  140;  Indels   19;  Gaps    8;

Qy         23 PTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLF---PYTT 75
              | | :|     ||    | |    :|:|  : |:|  : |:|| ||:  |: |   | |
Db        112 PNVNVF---VPPRDGFSGPAPRKSKLICEATNFTPKPITVSWLKDGKLVESGFTTDPVTI 168

Qy         76 RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCSDDEPRGVITYL 134
                 |    ||: : | : |::  |:: | ||| | | |  | :: |  |:      ::|:
Db        169 ENKGSTPQTYKVISTLTISEIDWLNLNVYTCRVDHRGLTFLKNVSSTCAASPSTDILTFT 228

Qy        135 IPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITS 194
              ||||  |:: : :  ||||| :| : | : ::|  :  : : :  :    | : | |
Db        229 IPPSFADIFLSKSANLTCLVSNLATYETLNISWASQSGEPLETKIKIMESHPNGTFSAKG 288

Qy        195 ILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGK--RLAPEVYMLPPSPEETG--T 250
              :  | :||   : : | | |   | |  : |:| | :  :  | ||:|||: |:
Db        289 VASVCVEDWNNRKEFVCTVTHRDLPSPQKKFISK-PNEVHKHPPAVYLLPPAREQLNLRE 347

Qy        251 TRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIW 310
              : |||||::|| |::|||||   :      : |: |  : |    :| :| : | :  |
Db        348 SATVTCLVKGFSPADISVQWLQRGQLLPQEKYVTSAPMPEPGAPGFYFTHSILTVTEEEW 407

Qy        311 EKGNLVTCRVVHEALP---GSRTLEKS 334
                |   || | |||||     ||::||
Db        408 NSGETYTCVVGHEALPHLVTERTVDKS 434


RESULT 12
GC4_HUMAN
ID   GC4_HUMAN      STANDARD;      PRT;   327 AA.
AC   P01861;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ig gamma-4 chain C region.
GN   IGHG4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83157104; PubMed=6299662;
RA   Ellison J.W., Buxbaum J.N., Hood L.E.;
RT   "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL   DNA 1:11-18(1981).
RN   [2]
RP   SEQUENCE OF 1-30 AND 81-326.
RX   MEDLINE=70207560; PubMed=4192699;
RA   Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
RT   "Human immunoglobulin subclasses. Partial amino acid sequence of the
RT   constant region of a gamma 4 chain.";
RL   Biochem. J. 117:33-47(1970).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K01316; AAB59394.1; ALT_INIT.
DR   PIR; A02150; G4HU.
DR   HSSP; P01842; 7FAB.
DR   Genew; HGNC:5528; IGHG4.
DR   MIM; 147130; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00407; IGC1; 2.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   Immunoglobulin domain; Immunoglobulin C region.
FT   NON_TER       1      1
FT   DOMAIN        1     98       CH1.
FT   DOMAIN       99    110       HINGE.
FT   DOMAIN      111    220       CH2.
FT   DOMAIN      221    327       CH3.
FT   DISULFID     14     14       INTERCHAIN (WITH A LIGHT CHAIN).
FT   DISULFID     27     83
FT   DISULFID    106    106       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    109    109       INTERCHAIN (WITH A HEAVY CHAIN).
FT   DISULFID    141    201
FT   DISULFID    247    305
```

```
SQ   SEQUENCE   327 AA;  35940 MW;  3EDBD811EF208E7A CRC64;

  Query Match             24.3%;  Score 451.5;  DB 1;  Length 327;
  Best Local Similarity   34.3%;  Pred. No. 3.9e-27;
  Matches  111;  Conservative   62;  Mismatches  120;  Indels   31;  Gaps   11;

Qy         35 RGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSE 90
              |  : ||  | |||  : |  | |:|     |  |    : ||        :    :|| |
Db         16 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSV 68

Qy         91 VNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEP-----------RGVITYLIPPSP 139
              |  :     : : ||||:| |  |  :   |  |   |             |   :| || |
Db         69 VTVPSSS-LGTKTYTCNVDHKPSNTKVDKRVESKYGPPCPSCPAPEFLGGPSVFLFPPKP 127

Qy        140 LD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILP 197
               | |  : ||::||:|:|:  |:  :   |  :  :   : ::   :  ::|  : |:|
Db        128 KDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLT 187

Qy        198 VDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTC 256
              |  :||: |: |:|:| :    |  | ::|:|  |:   |:|| |||| || |     ::||
Db        188 VLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLTC 247

Qy        257 LIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLV 316
              |::|||||:|:|:|  | : ::  :: || |  |  :| |||||||: |:|| |::||:
Db        248 LVKGFYPSDIAVEWESNGQPEN--NYKTTPPVLD--SDGSFFLYSRLTVDKSRWQEGNVF 303

Qy        317 TCRVVHEALPGSRTLEKSLHYSAG 340
              :| |:||||    | :|||  | |
Db        304 SCSVMHEALHNHYT-QKSLSLSLG 326


RESULT 13
MUC_SUNMU
ID   MUC_SUNMU      STANDARD;      PRT;   457 AA.
AC   P20768;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ig mu chain C region.
OS   Suncus murinus (House shrew) (Musk shrew).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Insectivora; Soricidae; Crocidurinae; Suncus.
OX   NCBI_TaxID=9378;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=89232144; PubMed=2497033;
RA   Ishiguro H., Ichihara Y., Namikawa T., Nagatsu T., Kurosawa Y.;
RT   "Nucleotide sequence of Suncus murinus immunoglobulin mu gene and
RT   comparison with mouse and human mu genes.";
RL   FEBS Lett. 247:317-322(1989).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13920; CAA32113.1; ALT_INIT.
DR   PIR; S03961; S03961.
DR   HSSP; P01842; 7FAB.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00407; IGc1; 3.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN        1    105       CH1.
FT   DOMAIN      106    220       CH2.
FT   DOMAIN      221    326       CH3.
FT   DOMAIN      327    457       CH4.
FT   DISULFID     13     13       INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT   DISULFID     27     89       BY SIMILARITY.
FT   DISULFID    136    200       BY SIMILARITY.
FT   DISULFID    217    217       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    248    307       BY SIMILARITY.
FT   DISULFID    295    295       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    355    417       BY SIMILARITY.
FT   DISULFID    456    456       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   CARBOHYD     45     45       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    212    212       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    276    276       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    283    283       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    444    444       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   457 AA;  50074 MW;  56C8C086DA4462E9 CRC64;

  Query Match             24.2%;  Score 449.5;  DB 1;  Length 457;
  Best Local Similarity   29.9%;  Pred. No. 8.3e-27;
  Matches  100;  Conservative   76;  Mismatches  138;  Indels   21;  Gaps    8;

Qy         21 IPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTT- 75
              :|| | :|     ||    |:  | ||:|  |||||  : ::||  |:  :      |:
Db        110 LPPNVSIF---VPPRNSFSGNHPRTSQLICQASGFSPRTIVMSWLQRGEPVQPSLVSTSA 166

Qy         76 ---RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFED--SSRRCSDDEPRGV 130
                 ||  |  || : | : ||: :|:|   :||   | |  |:   ||    ||   |:
Db        167 VEAEPKGSGPTTFRVISRLTITENEWLSQREFTCQALHKGLTFQKNVSSVCMGDDTSTGI 226

Qy        131 ITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATT 190
               :|:||:  :::   : :|||||  | : ::: ::| |:  ::: :     |  | ::|
Db        227 SVFLLPPTFANIFLTQSAQLTCLVTGLATYDSLDISWSRQNGEALQTHVNISESHPNSTF 286

Qy        191 SITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLA---PEVYMLPPSPEE 247
              :      |  ::|  || : | | |   | |: :|:::   | :|   | |::|||: |:
Db        287 TAKGHASVCREEWESGEKFTCTVQHSDLPSPLKQSLSR--PKDVANDPPSVFVLPPAQEQ 344

Qy        248 --TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLV 305
                : ::|||:: | | :: |||  : :     |: |: |  :      :|::| : |
Db        345 LKLRESASITCLVKDFSPPDVFVQWQHHGQPVDPKHYVTSNPTPEPQNPGLYFVHSILTV 404

Qy        306 NKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              ::  |  |   :| | ||||| | | ||::  ::|
Db        405 SEKDWSSGESFSCVVGHEALPLSVT-EKAVDKTSG 438


RESULT 14
GC1_HUMAN
ID   GC1_HUMAN      STANDARD;      PRT;   330 AA.
AC   P01857;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ig gamma-1 chain C region.
GN   IGHG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82274238; PubMed=6287432;
RA   Ellison J.W., Berson B.J., Hood L.E.;
RT   "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";
RL   Nucleic Acids Res. 10:4071-4079(1982).
RN   [2]
RP   SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX   MEDLINE=71064024; PubMed=5489771;
RA   Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA   Waxdal M.J., Edelman G.M.;
RT   "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
```

```
RT   acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL   Biochemistry 9:3161-3170(1970).
RN   [3]
RP   SEQUENCE OF 136-329 (EU).
RX   MEDLINE=71064025; PubMed=5530842;
RA   Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA   Edelman G.M.;
RT   "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RT   acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL   Biochemistry 9:3171-3181(1970).
RN   [4]
RP   SEQUENCE (MYELOMA PROTEIN NIE).
RX   MEDLINE=77070269; PubMed=826475;
RA   Ponstingl H., Hilschmann N.;
RT   "The rule of antibody structure. The primary structure of a
RT   monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The
RT   chymotryptic peptides of the H-chain, alignment of the tryptic
RT   peptides and discussion of the complete structure.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN   [5]
RP   SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX   MEDLINE=83289131; PubMed=6884994;
RA   Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT   "Three-dimensional structure determination of antibodies. Primary
RT   structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN   [6]
RP   DISULFIDE BONDS.
RX   MEDLINE=71064027; PubMed=4923144;
RA   Gall W.E., Edelman G.M.;
RT   "The covalent structure of a human gamma G-immunoglobulin. X.
RT   Intrachain disulfide bonds.";
RL   Biochemistry 9:3188-3196(1970).
RN   [7]
RP   DISULFIDE BONDS.
RX   MEDLINE=77070267; PubMed=1002129;
RA   Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT   "Rule of antibody structure. The primary structure of a monoclonal
RT   IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
RT   characterization of the protein, the L- and H-chains, the
RT   cyanogen bromide cleavage products, and the disulfide bridges.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=81208100; PubMed=7236608;
RA   Deisenhofer J.;
RT   "Crystallographic refinement and atomic models of a human Fc fragment
RT   and its complex with fragment B of protein A from Staphylococcus
RT   aureus at 2.9- and 2.8-A resolution.";
RL   Biochemistry 20:2361-2370(1981).
CC   -!- MISCELLANEOUS: NIE HAS THE G1M(17) ALLOTYPIC MARKER, 97-K, & THE
CC       G1M(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE G1M(3)
CC       MARKER & THE G1M (NON-1) MARKERS.
CC   -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
CC       35,116,198,269 & 272.
CC   -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
CC       155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
CC       268-272.
CC   -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
CC       RESIDUES 198,267&272.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J00228; AAC82527.1; ALT_INIT.
DR   PIR; A02146; GHHU.
DR   PDB; 1FC1; 15-JUL-92.
DR   PDB; 1FC2; 15-JUL-92.
DR   Genew; HGNC:5525; IGHG1.
DR   MIM; 147100; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW   3D-structure.
FT   NON_TER       1      1
FT   DOMAIN        1     98       CH1.
FT   DOMAIN       99    110       HINGE.
FT   DOMAIN      111    223       CH2.
FT   DOMAIN      224    330       CH3.
FT   DISULFID     27     83
FT   DISULFID    103    103       INTERCHAIN (WITH LIGHT CHAIN).
FT   DISULFID    109    109       INTERCHAIN (WITH HEAVY CHAIN).
FT   DISULFID    112    112       INTERCHAIN (WITH HEAVY CHAIN).
FT   DISULFID    144    204
FT   DISULFID    250    308
FT   CARBOHYD    180    180       N-LINKED (GLCNAC. . .).
FT   MOD_RES     330    330       REMOVED POST-TRANSLATIONALLY.
FT   VARIANT      97     97       K -> R (IN G1M(3) MARKER).
FT                                /FTId=VAR_003886.
FT   VARIANT     239    239       D -> E (IN G1M(NON-1) MARKER).
FT                                /FTId=VAR_003887.
FT   VARIANT     241    241       L -> M (IN G1M(NON-1) MARKER).
FT                                /FTId=VAR_003888.
FT   STRAND      123    126
FT   HELIX       130    134
FT   TURN        136    137
FT   STRAND      141    148
FT   STRAND      158    162
FT   TURN        163    164
FT   STRAND      165    166
FT   STRAND      175    178
FT   STRAND      183    190
FT   HELIX       193    197
FT   TURN        198    199
FT   STRAND      202    206
FT   STRAND      215    219
FT   STRAND      227    227
FT   STRAND      230    234
FT   HELIX       238    240
FT   TURN        241    242
FT   STRAND      245    256
FT   STRAND      260    266
FT   TURN        267    268
FT   STRAND      269    270
FT   STRAND      274    276
FT   STRAND      280    281
FT   TURN        283    284
FT   STRAND      287    296
FT   HELIX       297    301
FT   TURN        302    303
FT   STRAND      306    312
FT   TURN        313    314
FT   TURN        316    317
FT   STRAND      320    324
SQ   SEQUENCE   330 AA;  36106 MW;  3770EE106C2FA33D CRC64;

  Query Match             23.7%;  Score 440.5;  DB 1;  Length 330;
  Best Local Similarity   32.6%;  Pred. No. 2.7e-26;
  Matches  111;  Conservative   67;  Mismatches  125;  Indels   37;  Gaps   13;

Qy         22 PPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRP 77
              |    | ||    |    | | |||  : |  | |:|    |  |    : ||
Db          6 PSVFPLAPSSKSTSG---GTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL-- 57

Qy         78 KREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE----------DSSRRC----S 123
              :    :|| | | :     : : || |:| |  |  :          | :  |    :
```

```
Db         58 --QSSGLYSLSSVVTVPSSS-LGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPA 114

Qy        124 DDEPRGVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQR 181

Db        115 PELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 174

Qy        182 STKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYML 241

Db        175 REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 234

Qy        242 PPSPEE-TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300

Db        235 PPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN--NYKTTPPVLD--SDGSFFLY 290

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340

Db        291 SKLTVDKSRWQQGNVFSCSVMHEALHNHYT-QKSLSLSPG 329


RESULT 15
MUC_MESAU
ID   MUC_MESAU      STANDARD;      PRT;   454 AA.
AC   P06337;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig mu chain C region.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85297761; PubMed=2994005;
RA   McGuire K.L., Duncan W.R., Tucker P.W.;
RT   "Phylogenetic conservation of immunoglobulin heavy chains: direct
RT   comparison of hamster and mouse Cmu genes.";
RL   Nucleic Acids Res. 13:5611-5628(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02804; CAA26574.1; -.
DR   PIR; A02168; MHHY.
DR   HSSP; P01854; 1IGE.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 2.
DR   SMART; SM00407; IGc1; 2.
DR   PROSITE; PS00290; IG_MHC; 3.
KW   Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN        1    105       CH1.
FT   DOMAIN      106    218       CH2.
FT   DOMAIN      219    324       CH3.
FT   DOMAIN      325    454       CH4.
FT   DISULFID     13     13       INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT   DISULFID     27     88       BY SIMILARITY.
FT   DISULFID    135    198       BY SIMILARITY.
FT   DISULFID    215    215       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    245    304       BY SIMILARITY.
FT   DISULFID    292    292       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   DISULFID    352    414       BY SIMILARITY.
FT   DISULFID    453    453       INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT   CARBOHYD     45     45       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    210    210       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    238    238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   454 AA;  50196 MW;  4DA2134612BE1469 CRC64;

  Query Match           23.3%;  Score 433.5;  DB 1;  Length 454;
  Best Local Similarity   30.7%;  Pred. No. 1.3e-25;
  Matches  106;  Conservative   68;  Mismatches  154;  Indels   17;  Gaps    9;

Qy          5 HHHHT---LSLPESGPVTIIPPTVKLFHSSCDP-RGDAHSTIQLLCLVSGFSPAKVHVTW 60

Db         91 HHGNTNKDLRVP-IPVVTEMNPNVSVFVPSRDAFSGPAPRKSRLFCEASNFSPKQITVSW 149

Qy         61 LVDGQEAENLF---PYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-E 116

Db        150 LRDGKPVKSGFTTEPVTPEDRGSGPRTYKVISTLTITESDWLNLSVYTCRVDHRGLTFWK 209

Qy        117 DSSRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIG 176

Db        210 NVSSTCAASPSTDIQAFPIPPSFVGIFLNKSATLTCLVTNLATYDTLNISWSSRSGEPLE 269

Qy        177 SASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGK--RL 234

Db        270 TKTKLTESHPNGTFSAIGEANVCVEDWDSGKEFVCTVTHRDLPSPQKKFISK-PREMNKT 328

Qy        235 APEVYMLPPSPEE--TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHG 292

Db        329 PPAVYQQPLAREQLILRESATVTCLVKGFSPADIFVQWLQRGQPLSQDKYVTSAPMREPQ 388

Qy        293 TDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALP---GSRTLEKS 334

Db        389 APHLYFTHSVLTVTEEEWNSGETYTCVVGHEALPHMVTERTVDRS 433


Search completed: July  9, 2003, 13:40:01
Job time : 12 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 19 Seconds
(without alignments)
1725.360 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHHTLSLPESGPVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1051	56.6	429	1	EHRT	Ig epsilon chain C
2	940.5	50.6	388	1	EHMS	Ig epsilon chain C
3	871.5	46.9	548	2	S38864	Ig epsilon chain C
4	862.5	46.4	423	1	EHMSS	Ig epsilon chain C
5	745	40.1	428	1	EHHU	Ig epsilon chain C
6	741	39.9	426	2	I36948	Ig epsilon-chain -
7	587	31.6	572	2	B46529	Ig Y heavy chain (
8	560	30.1	504	2	S00390	Ig gamma chain (cl
9	499	26.9	627	2	S14683	Ig mu chain precur
10	490	26.4	479	1	MHRBM	Ig mu chain C regi
11	488.5	26.3	453	2	S37768	Ig mu chain C regi
12	488.5	26.3	474	2	S15590	Ig heavy chain - h
13	488	26.3	458	1	MHRB	Ig mu chain C regi
14	484.5	26.1	391	1	MHHUBT	Ig mu heavy chain
15	477	25.7	452	1	MHHU	Ig mu chain C regi
16	477	25.7	473	1	MHHUM	Ig mu chain C regi
17	469.5	25.3	433	2	S31436	Ig upsilon chain -
18	468	25.2	326	1	G2HU	Ig gamma-2 chain C
19	464.5	25.0	450	1	MHDG	Ig mu chain C regi
20	464	25.0	328	2	I47161	Ig gamma 3 chain c
21	461.5	24.8	448	2	S03186	Ig heavy chain C r
22	459.5	24.7	328	2	I47160	Ig gamma 2b chain
23	456.5	24.6	328	2	I47159	Ig gamma 2a chain
24	455.5	24.5	455	1	MHMS	Ig mu chain C regi
25	455.5	24.5	455	2	A24976	Ig mu chain C regi
26	455.5	24.5	476	1	MHMSM	Ig mu chain C regi
27	454	24.4	328	2	I47158	Ig gamma 1 chain c
28	454	24.4	592	2	S25705	Ig mu chain - shee
29	451.5	24.3	327	1	G4HU	Ig gamma-4 chain C
30	451.5	24.3	472	2	S31459	Ig gamma-1 chain -
31	449.5	24.2	457	2	S03961	Ig mu chain C regi
32	448	24.1	343	2	S25644	Ig mu chain C regi
33	447.5	24.1	454	1	MHHY	Ig mu chain C regi
34	441	23.7	107	2	I68726	IgE chain C3 regio
35	440.5	23.7	330	1	GHHU	Ig gamma-1 chain C
36	439	23.6	107	2	I68730	IgE chain C3 regio
37	439	23.6	374	2	S69339	Ig heavy chain V r
38	437.5	23.5	277	2	I47162	Ig gamma 4 chain c
39	437.5	23.5	377	2	A60764	Ig gamma-3 chain C
40	437.5	23.5	377	2	A23511	Ig gamma-3 chain C
41	432	23.3	444	2	PC4436	monoclonal antibod
42	430.5	23.2	470	2	S22080	Ig heavy chain pre
43	429.5	23.1	308	2	C30554	Ig heavy chain C r
44	426	22.9	322	2	PS0019	Ig gamma-2a chain
45	424.5	22.8	329	2	S00847	Ig gamma-2c chain

ALIGNMENTS

RESULT 1
EHRT
Ig epsilon chain C region - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
C;Accession: A93442; A90937; A02143
R;Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A;Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A;Reference number: A93442; MUID:83064537; PMID:6292865
A;Accession: A93442
A;Molecule type: mRNA
A;Residues: 1-429 <HEL>
A;Experimental source: strain LOU/c/Wsl, immunocytoma IR2
R;Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A;Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction, i
A;Reference number: A90937; MUID:83182019; PMID:6820340
A;Contents: myeloma IR162
A;Accession: A90937
A;Molecule type: mRNA
A;Residues: 'N',169-307,'L',309-342 <KIN>
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;19-80/Domain: immunoglobulin homology <IM1>
F;118-186/Domain: immunoglobulin homology <IM2>
F;223-291/Domain: immunoglobulin homology <IM3>
F;327-398/Domain: immunoglobulin homology <IM4>
F;46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predicte

Query Match 56.6%; Score 1051; DB 1; Length 429;
Best Local Similarity 62.6%; Pred. No. 3.7e-70;
Matches 206; Conservative 35; Mismatches 78; Indels 10; Gaps 3;

```
Qy         17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTR 76
              || |  ||| | ||||||    |||||| | | |   | : ||:| ::      | |
Db         97 PVNITKPTVDLLHSSCDPNA-FHSTIQLYCFVYGHIQNDVSIHWLMDDRKI-----YETH 150

Qy         77 PK----REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVIT 132
                :    :| |:  |  | :|||| |||| :|:|| |   |  :   :|||||||||||||
Db        151 AQNVLIKEEGKLASTYSRLNITQQQWMSESTFTCKVTSQGENYWAHTRRCSDDEPRGVIT 210

Qy        133 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSI 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        211 YLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHNATTSI 270

Qy        193 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR 252
              |||||||||||||||||||||||||||||||| |||| |||||:  |  ||    |
Db        271 TSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKAPGKRSAPEVYVFLPPEEEEKDKR 330
```

```
Qy        253 TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEK 312
              |:||||: |:| :|||||| :::      |:|| | | :|::  ||::||: | |::| :
Db        331 TLTCLIQNFFPEDISVQWLQDSKLIPKSQHSTTTPLKYNGSNQRFFIFSRLEVTKALWTQ 390

Qy        313 GNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
                  ||||:||||   | ||:::  | ||
Db        391 TKQFTCRVIHEALREPRKLERTISKSLGN 419


RESULT 2
EHMS
Ig epsilon chain C region (version 1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C;Accession: A02144
R;Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.
Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982
A;Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.
A;Reference number: A02144; MUID:83117774; PMID:6818553
A;Accession: A02144
A;Molecule type: mRNA
A;Residues: 1-388 <LIU>
A;Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;1-44/Domain: immunoglobulin homology (fragment) <IM1>
F;81-149/Domain: immunoglobulin homology <IM2>
F;186-254/Domain: immunoglobulin homology <IM3>
F;290-361/Domain: immunoglobulin homology <IM4>
F;10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status predi

  Query Match             50.6%;  Score 940.5;  DB 1;  Length 388;
  Best Local Similarity   55.2%;  Pred. No. 4.7e-62;
  Matches  191;  Conservative   46;  Mismatches  100;  Indels    9;  Gaps    5;

Qy          2 FHHHHHHTLSLPES-----GPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKV 56
              |  |  |  |  ||      || |  ||::| ||||||    |||||| | : |      |
Db         40 FTCHVTHPPSFNESRTILVRPVNITEPTLELLHSSCDPNA-FHSTIQLYCFIYGHILNDV 98

Qy         57 HVTWLVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
              |:||:| :|  :    |   | | |:  |  |::|||: |||| :|:|| |   |  :
Db         99 SVSWLMDDREITDTLAQTVLIKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYL 157

Qy        117 DSSRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIG 176
                :||| | |||||||||||||||||:|| |||||||:|||||:|: ||| :|:| |:
Db        158 AHTRRCPDHEPRGVITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV- 216

Qy        177 SASQRSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLA 235
              ||||  |||| :|||||||||| |||||| |||| |||| |||||||||| ||:| |
Db        217 SASQWYTKHHNNATTSITSILPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSA 276

Qy        236 PEVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDP 295
              ||||: ||  ||:   ||:||||: |:| :|||||| : :      |:|| | | :|::
Db        277 PEVYVFPPPEEESEDKRTLTCLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNQ 336

Qy        296 SFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              ||::||: | |::| :    ||:|:|||   | |||::  | ||
Db        337 GFFIFSRLEVAKTLWTQRKQFTCQVIHEALQKPRKLEKTISTSLGN 382


RESULT 3
S38864
Ig epsilon chain C region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C;Accession: S38864
R;Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A;Description: Combination of a defined specificity and desired isotype by cloning of an
A;Reference number: S38864
A;Accession: S38864
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-548 <KIP>
A;Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;353-421/Domain: immunoglobulin homology <IMM>

  Query Match             46.9%;  Score 871.5;  DB 2;  Length 548;
  Best Local Similarity   53.9%;  Pred. No. 9e-57;
  Matches  178;  Conservative   45;  Mismatches   92;  Indels   15;  Gaps    5;

Qy         12 LPESGPVTIIPPTVKLFHSSCDPRGD--AHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              :||   | | ||  |           |    ||||| | : |      | |:||:| :|  :
Db        229 VPEVSSVFIFPPKPK----------DVLTRSTIQLYCFIYGHILNDVSVSWLMDDREITD 278

Qy         70 LFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRG 129
                  | | |:  |  |::|||: |||| :|:|| |   |  :   :||| | ||||
Db        279 TLAQTVLIKEE-GKLASTCSKLNITEQQWMSESTFTCKVTSQGVDYLAHTRRCPDHEPRG 337

Qy        130 VITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHH-HA 188
              |||||||||||||:|| |||||||:|||||:|: ||| :|:| |: ||||  |||| :|
Db        338 VITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEKKTSV-SASQWYTKHHNNA 396

Qy        189 TTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEET 248
              |||||||||| ||||||| |||| |||| |||||||||| ||:| |||||: ||  ||:
Db        397 TTSITSILPVVAKDWIEGYGYQCIVDHPDFPKPIVRSITKTPGQRSAPEVYVFPPPEEES 456

Qy        249 GTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKS 308
                ||:|||: |:| :|||||| : :      |:|| | | :|::  ||::||: | |:
Db        457 EDKRTLTCLIQNFFPEDISVQWLGDGKLISNSQHSTTTPLKSNGSNRGFFIFSRLEVAKT 516

Qy        309 IWEKGNLVTCRVVHEALPGSRTLEKSLHYS 338
              :| :    ||:|:||||   | |||::  |
Db        517 LWTQRKQFTCQVIHEALQKPRKLEKTISTS 546


RESULT 4
EHMSS
Ig epsilon chain C region (version 2) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
C;Accession: A02145
R;Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.
EMBO J. 1, 1117-1123, 1982
A;Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison wi
A;Reference number: A90966; MUID:84236092; PMID:6329728
A;Accession: A02145
A;Molecule type: DNA
A;Residues: 1-423 <ISH>
A;Note: the sequence was determined from the germline gene
C;Genetics:
A;Introns: 91/1; 199/1; 307/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;16-77/Domain: immunoglobulin homology <IMM1>
F;115-183/Domain: immunoglobulin homology <IMM2>
F;220-288/Domain: immunoglobulin homology <IMM3>
F;325-396/Domain: immunoglobulin homology <IMM4>
F;23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
F;43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             46.4%;  Score 862.5;  DB 1;  Length 423;
  Best Local Similarity   51.8%;  Pred. No. 3e-56;
  Matches  183;  Conservative   46;  Mismatches  103;  Indels   21;  Gaps    8;

Qy          2 FHHHHHHTLSLPESGPVTIIPPTVKLFHS----------SCDPRGDAHSTIQLLCLVSGF 51
              |  |  |  |  ||  : : | |    ||           |||    |||||| | : |
Db         73 FTCHVTHPPSFNESRTILVRPVT----HSLSPPWSYSIHRCDPNA-FHSTIQLYCFIYGH 127

Qy         52 SPAKVHVTWLVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHN 111
                 | |:||:| :|  :    |   | | |:  |  |::|||: |||| :|:|| |
```

```
Db        128 ILNDVSVSWLMDDREITDTLAQTVLIKEE-GKLASTCSKLNITEQQWMSESTFTCRVTSQ 186

Qy        112 GSIFEDSSRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRER 171
              |  :   :||| | |||| ||||||||||||:|| |||||||:|||||:|: ||| :|:
Db        187 GVDYLAHTRRCPDHEPRGAITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEK 246

Qy        172 KKSIGSASQRSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLP 230
              | |: ||||  |||| :|||||||||| ||||||| |||| || | ||||||||| ||
Db        247 KTSV-SASQWYTKHHNNATTSITSILPVVAKDWIEGYGYQCVVDRPDFPKPIVRSIT-LP 304

Qy        231 --GKRLAPEVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQ 288
                :| |||||: ||  ||:   ||:|||: |:| :|||||| : :       |:|| |
Db        305 QVSQRSAPEVYVFPPPEEESEDKRTLTCLIQNFFPEDISVQWLGDGKLISNSQHSTTTPL 364

Qy        289 KDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341
              | :|::  ||::||: | |::| :    ||:|:||||   | |||::  | ||
Db        365 KSNGSNQGFFIFSRLEVAKTLWTQRKQFTCQVIHEALQKPRKLEKTISTSLGN 417


RESULT 5
EHHU
Ig epsilon chain C region - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1981 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C;Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C46
R;Flanagan, J.G.; Rabbitts, T.H.
EMBO J. 1, 655-660, 1982
A;Title: The sequence of a human immunoglobulin epsilon heavy chain constant region gene
A;Reference number: A22771; MUID:84236029; PMID:6234164
A;Accession: A22771
A;Molecule type: DNA
A;Residues: 1-428 <FLA>
A;Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035
R;Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.
EMBO J. 1, 1539-1544, 1982
A;Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudog
A;Reference number: A23195; MUID:84207910; PMID:6327276
A;Accession: A23195
A;Molecule type: DNA
A;Residues: 2-428 <UED>
A;Cross-references: GB:J00222; NID:g184755
R;Zhang, K.; Saxon, A.; Max, E.E.
J. Exp. Med. 176, 233-243, 1992
A;Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing
A;Reference number: PH1214; MUID:92308839; PMID:1613458
A;Accession: PH1214
A;Molecule type: DNA
A;Residues: 320-428 <ZHA>
A;Cross-references: EMBL:X63693; GB:S38668; NID:g32987
R;Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Kikuchi, M.; Sugi
Nucleic Acids Res. 11, 719-726, 1983
A;Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon cha
A;Reference number: A93491; MUID:83168897; PMID:6300763
A;Accession: A93491
A;Molecule type: mRNA
A;Residues: 1-428 <SEN>
A;Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035
R;Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.
Cell 29, 691-699, 1982
A;Title: Duplication and deletion in the human immunoglobulin epsilon genes.
A;Reference number: A90824; MUID:83001945; PMID:6288268
A;Accession: A90824
A;Molecule type: DNA
A;Residues: 1-358,'L',360-428 <MAX>
A;Cross-references: GB:J00222; NID:g184755
A;Note: this sequence difference may be due to polymorphism
R;Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-3
A;Reference number: A94418
A;Accession: A94418
A;Molecule type: protein
A;Residues: 'GAWTL',6,'X',8-16,'B',18-43,'B',45-52,55-92,95-97,'B',99-121,'B',123,'L',12
A;Experimental source: myeloma protein Nd
R;Kenten, J.H.; Molgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.; G
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A;Title: Cloning and sequence determination of the gene for the human immunoglobulin eps
A;Reference number: A93933; MUID:83065234; PMID:6815656
A;Accession: B93933
A;Molecule type: mRNA
A;Residues: 1-40;68-114;427-428 <KEN>
A;Cross-references: GB:L00022; NID:g185035
R;Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A;Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragmen
A;Reference number: S02438; MUID:88083554; PMID:3121387
A;Accession: S02438
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 98-352 <IKE>
R;Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A;Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces
A;Reference number: A53116; MUID:94103254; PMID:8276835
A;Accession: A53116
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 320-428 <ZH2>
A;Experimental source: myeloma U266-derived cell line AF-10
A;Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBIP:141702)
R;Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A;Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of
A;Reference number: A46536; MUID:93122085; PMID:8419166
A;Accession: C46536
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 382-426 <HEL>
A;Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A;Experimental source: B cell myeloma U-266
A;Note: sequence extracted from NCBI backbone (NCBIP:125297)
A;Accession: D46536
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 382-391 <HE2>
A;Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A;Experimental source: B cell myeloma U-266
A;Note: sequence extracted from NCBI backbone (NCBIP:125299)
A;Accession: A46536
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 401-428 <HE3>
A;Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A;Experimental source: B cell myeloma U-266
A;Note: sequence extracted from NCBI backbone (NCBIP:123483)
C;Genetics:
A;Gene: GDB:IGHE
A;Cross-references: GDB:119335; OMIM:147180
A;Map position: 14q32.33-14q32.33
A;Introns: 1/1; 104/1; 211/1; 319/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F;22-87/Domain: immunoglobulin homology <IM1>
F;128-195/Domain: immunoglobulin homology <IM2>
F;232-301/Domain: immunoglobulin homology <IM3>
F;338-407/Domain: immunoglobulin homology <IM4>
F;14/Disulfide bonds: interchain (to light chain) #status predicted
F;15-105,29-85,135-193,239-299,345-405/Disulfide bonds: #status predicted
F;21,49,99,146,252,275/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;121,209/Disulfide bonds: interchain (to heavy chain) #status predicted

  Query Match            40.1%;  Score 745;  DB 1;  Length 428;
  Best Local Similarity  44.4%;  Pred. No. 1.5e-47;
  Matches 143;  Conservative   62;  Mismatches 109;  Indels    8;  Gaps    5;
```

```
Qy         22 PPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRPKREG 81
              |||||:  ||||  |    ||||||||||::|  :::||| |||  :     :|    :
Db        111 PPTVKILQSSCDGGGHFPPTIQLLCLVSGYTPGTINITWLEDGQVMD--VDLSTASTTQE 168

Qy         82 GQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIPPSPLD 141
              |:  | |||: ::|  |:|  |||| | : |  ||||:::|:|  ||||  ||  ||| |
Db        169 GELASTQSELTLSQKHWLSDRTYTCQVTYQGHTFEDSTKKCADSNPRGVSAYLSRPSPFD 228

Qy        142 LYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSASQRSTKHHHATTSITSILPVDA 200
              |:   :| :||||:||  |:  : :|| |   | :  ::::  |  : | ::|| |||
Db        229 LFIRKSPTITCLVVDLAPSKGTVNLTWSRASGKPVNHSTRKEEKQRNGTLTVTSTLPVGT 288

Qy        201 KDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGT--TRTVTCLI 258
              :|||||| ||||| ||| |: ::|| ||  | | |||||    :|| |:   ||: |||
Db        289 RDWIEGETYQCRVTHPHLPRALMRSTTKTSGPRAAPEVYAF-ATPEWPGSRDKRTLACLI 347

Qy        259 RGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVTC 318
              : | | :|||||| | :     |:||:|:|  |:   ||::||: | :: ||: :   |
Db        348 QNFMPEDISVQWLHNEVQLPDARHSTTQPRKTKGS--GFFVFSRLEVTRAEWEQKDEFIC 405

Qy        319 RVVHEALPGSRTLEKSLHYSAG 340
              | ||||   |:|:::::  : |
Db        406 RAVHEAASPSQTVQRAVSVNPG 427


RESULT 6
I36948
Ig epsilon-chain - chimpanzee (fragment)
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I36948
R;Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A;Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and oranguta
A;Reference number: I36948; MUID:87147196; PMID:3103123
A;Accession: I36948
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-426 <RES>
A;Cross-references: GB:M15398; NID:g176797; PIDN:AAA35416.1; PID:g176798
C;Genetics:
A;Introns: 103/1; 209/1; 317/1
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;336-405/Domain: immunoglobulin homology <IMM>

  Query Match             39.9%;  Score 741;  DB 2;  Length 426;
  Best Local Similarity   44.5%;  Pred. No. 2.9e-47;
  Matches  143;  Conservative   60;  Mismatches  110;  Indels    8;  Gaps    5;

Qy         23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRPKREGG 82
              ||||:  ||||  |    ||||||||||::|  :::||| |||  :     :|    : |
Db        110 PTVKVLQSSCDGGGHFPPTIQLLCLVSGYTPGTINITWLEDGQVMD--VDLSTASATQEG 167

Qy         83 QTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIPPSPLDL 142
              :  | |||: ::|  |:|  |||| | : |  ||||:::|:|  ||||  ||  ||| ||
Db        168 ELASTQSELTLSQKHWLSDRTYTCQVTYQGGTFEDSTKKCADSNPRGVSAYLSRPSPFDL 227

Qy        143 YENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSASQRSTKHHHATTSITSILPVDAK 201
              :   :| :||||:||  |:  : :|| |   | :  ::::  |  : | ::|| |||  :
Db        228 FIRKSPTITCLVVDLAPSKGTVNLTWSRASGKPVNHSTRKQEKQRNGTLTVTSTLPVGTR 287

Qy        202 DWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGT--TRTVTCLIR 259
              |||||| ||||| ||| |: :||| ||  | | |||||    :|| |:   ||: |||:
Db        288 DWIEGETYQCRVTHPHLPRALVRSTTKTSGPRAAPEVYAF-ATPEGPGSRDKRTLACLIQ 346

Qy        260 GFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVTCR 319
              | | :|||||| | :     |:||:| |  |:   ||::||: | :: ||: :   ||
Db        347 NFMPEDISVQWLHNEVQLPDARHSTTQPHKTKGS--GFFVFSRLEVTRAEWEQKDEFICR 404

Qy        320 VVHEALPGSRTLEKSLHYSAG 340
              ||||   |:|:::::  : |
Db        405 AVHEAASPSQTVQRTVSVNPG 425


RESULT 7
B46529
Ig Y heavy chain (7.8S) - duck
N;Alternate names: Ig gamma chain (7.8S)
C;Species: Anas platyrhynchos (domestic duck)
C;Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: B46529; S20759
R;Magor, K.E.; Warr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.
J. Immunol. 149, 2627-2633, 1992
A;Title: Structural relationship between the two IgY of the duck, Anas platyrhynchos: mo
A;Reference number: A46529; MUID:93017865; PMID:1401901
A;Accession: B46529
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-572 <MAG>
A;Cross-references: EMBL:X65219; NID:g62442; PIDN:CAA46322.1; PID:g62443
A;Experimental source: spleen
A;Note: sequence extracted from NCBI backbone (NCBIP:116127)
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;37-120/Domain: immunoglobulin homology <IMM>

  Query Match             31.6%;  Score 587;  DB 2;  Length 572;
  Best Local Similarity   37.3%;  Pred. No. 9.8e-36;
  Matches  121;  Conservative   60;  Mismatches  129;  Indels   14;  Gaps    8;

Qy         22 PPTVKLFHSS-CDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRPKRE 80
              || |:: ||| |   ||   :::|||:::|||| | | |||||  | :|    |||:||
Db        249 PPEVQVLHSSVCSTLGD--DSVELLCVITGFSPPPVEVEWLVDGAPA-HLVATMTRPQRE 305

Qy         81 -GGQTFSLQSEVNITQGQWMSSNTYTCHVKH--NGSIFEDSSRRCSDDEPRG---VITYL 134
               | :|:   |: |:::  | :   :|| |||   |   :  :| |     :   :  ::
Db        306 AGSKTYMATSQTNVSREDWKAGKAFTCRVKHPATGGTAQGHARFCPGSGAQSCSPIQIFV 365

Qy        135 IPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITS 194
              :||||  ||     |: |||::| |: :::::| ||:  ::        |:| : | : :|
Db        366 VPPSPGSLYIRQDAKVHCLVVNLPSDASLSISWTREKSGALRPDPMVLTEHFNGTFTASS 425

Qy        195 ILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR-T 253
              | :  :||: || : | | |   | |: :|| |  ||  || ::  ||  ||      |
Db        426 SLAISTQDWLAGERFTCTVQHEDLPVPLGKSIAKHAGKVTAPYIFTFPPHAEELSLAEVT 485

Qy        254 VTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKG 313
              :|||:||| |  : |||| |:          || | |:   | :||||||:| | |: |: |
Db        486 LTCLVRGFQPEHVEVQWLRNHNSVPAAEFVTTPPLKEPNGDGTFFLYSKMTVPKASWQGG 545

Qy        314 NLVTCRVVHEALP---GSRTLEKS 334
                 | |||| ||     | |:|:
Db        546 VSYACMVVHEGLPMRFTQRPLQKT 569


RESULT 8
S00390
Ig gamma chain (clone 36) - chicken (fragment)
N;Alternate names: Ig nu chain
C;Species: Gallus gallus (chicken)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C;Accession: S00390
R;Parvari, R.; Avivi, A.; Lentner, F.; Ziv, E.; Tel-Or, S.; Burstein, Y.; Schechter, I.
EMBO J. 7, 739-744, 1988
A;Title: Chicken immunoglobulin gamma-heavy chains: limited VH gene repertoire, combinat
A;Reference number: S00390; MUID:88283642; PMID:3135182
A;Accession: S00390
A;Molecule type: mRNA
A;Residues: 1-504 <PAR>
A;Cross-references: EMBL:X07174
A;Note: this sequence was determined from the differentiated gene
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin

  Query Match             30.1%;  Score 560;  DB 2;  Length 504;
  Best Local Similarity   35.7%;  Pred. No. 8.2e-34;
```

```
  Matches  119;  Conservative   57;  Mismatches  137;  Indels   20;  Gaps    8;

Qy          17 PVTIIPPTVKLFH-SSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTT 75
               ||  : | |:: | ||| |      :::|||||:|||||   | |||||     |   :
Db         172 PVPPVAPEVQVLHASSCTP--SQSESVELLCLVTGFSPASAEVEWLVDG--VGGLLVASQ 227

Qy          76 RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKH--NGSIFEDSSRRCSD--DEPRGVI 131
               |    | |:|| | ||::    |    :|:| |:|    :: ||  : | |      :
Db         228 SPAVRSGSTYSLSSRVNVSGTDWREGKSYSCRVRHPATNTVVEDHVKGCPDGAQSCSPIQ 287

Qy         132 TYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTS 191
               | ||||| :|| :   || |||::| |: :::||| ||:  ::         :| : | |
Db         288 LYAIPPSPGELYISLDAKLRCLVVNLPSDSSLSVTWTREKSGNLRPDPMVLQEHFNGTYS 347

Qy         192 ITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTT 251
               :| :||  :||: || : | | |   | |: :|: :  |    | :|  ' ||| :
Db         348 ASSAVPVSTQDWLSGERFTCTVQHEELPLPLSKSVYRNTGPTTPPLIYPFAPHPEELSLS 407

Qy         252 R-TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDH-------GTDPSFFLYSRM 303
               | |::||:||| | :| ::|| ::          ||    :         |   :||:||:|
Db         408 RVTLSCLVRGFRPRDIEIRWLRDHRAVPATEFVTTAVLPEERTANGAGGDGDTFFVYSKM 467

Qy         304 LVNKSIWEKGNLVTCRVVHEALP---GSRTLEK 333
               |  : |  | :  |  ||||||     |||:|
Db         468 SVETAKWNGGTVFACMAVHEALPMRFSQRTLQK 500


RESULT 9
S14683
Ig mu chain precursor, membrane-bound (clone 201) - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C;Accession: S14683; S08047
R;Friedlander, R.M.; Nussenzweig, M.C.; Leder, P.
Nucleic Acids Res. 18, 4278, 1990
A;Title: Complete nucleotide sequence of the membrane form of the human IgM heavy chain.
A;Reference number: S14683; MUID:90332450; PMID:2115996
A;Accession: S14683
A;Molecule type: mRNA
A;Residues: 1-627 <FRI>
A;Cross-references: EMBL:X17115; NID:g33450; PIDN:CAA34971.1; PID:g33451
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin; membrane protein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-627/Product: Ig mu chain #status predicted <MAT>
F;34-117/Domain: immunoglobulin homology <IMM>

  Query Match             26.9%;  Score 499;  DB 2;  Length 627;
  Best Local Similarity   31.5%;  Pred. No. 3.5e-29;
  Matches  106;  Conservative   72;  Mismatches  135;  Indels   24;  Gaps    9;

Qy          17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
               |: :|   || | :|     ||    |:  |  :|:|  :|||| :: |:|| :|::  :
Db         254 PLPVIAELPPKVSVF---VPPRDGFFGNPRSKSKLICQATGFSPRQIQVSWLREGKQVGS 310

Qy          70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                ||        |  |  |: : | : | :  :| : :|| | | |  | :::|  |
Db         311 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 368

Qy         124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
               |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db         369 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 428

Qy         184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
               | :|| |       :   ||  || : | | |    | |: ::|::  |   |  |:||:||
Db         429 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 488

Qy         243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
               |: |:     : |:|||: || |::: |||:   :       : |: |  :       :| :
Db         489 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 548

Qy         301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
               | : |::  |  |   || | |||||     ||::||
Db         549 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 585


RESULT 10
MHRBM
Ig mu chain C region, membrane-bound form - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Dec-1986 #sequence_revision 30-Jun-1991 #text_change 23-Aug-1997
C;Accession: A02165; A02164
R;Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.
J. Immunol. 132, 490-495, 1984
A;Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-Ha2 a
A;Reference number: A02164; MUID:84088930; PMID:6418803
A;Contents: a2 allotype
A;Accession: A02165
A;Molecule type: mRNA
A;Residues: 439-479 <BE2>
A;Accession: A02164
A;Molecule type: mRNA
A;Residues: 1-438,'GKPTLYNVSLIMSDTASTCY' <BER>
A;Note: the sequence of residues 1-438 was assumed to be identical with the correspondin
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F;21-92/Domain: immunoglobulin homology <IMM1>
F;130-202/Domain: immunoglobulin homology <IMM2>
F;242-310/Domain: immunoglobulin homology <IMM3>
F;349-420/Domain: immunoglobulin homology <IMM4>
F;439-479/Domain: carboxyl-terminal <CTS>
F;14/Disulfide bonds: interchain (to light chain) #status predicted
F;28-90,137-200,249-308,356-418/Disulfide bonds: #status predicted
F;46,114,212,261,277,284/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;219/Disulfide bonds: interchain (to heavy chain) #status predicted
F;296/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

  Query Match             26.4%;  Score 490;  DB 1;  Length 479;
  Best Local Similarity   32.4%;  Pred. No. 1.1e-28;
  Matches  115;  Conservative   62;  Mismatches  154;  Indels   24;  Gaps    9;

Qy           5 HHHHTLSLPESGPV-TIIPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVT 59
               | :    |  | || : :|| | :|      ||     |       :|:|  :|||| :: |:
Db          94 HSNSNRDLRVSFPVDSELPPNVSVF---IPPRDSFSGSGTRKSRLICQATGFSPKQISVS 150

Qy          60 WLVDGQEAEN---LFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
               || |||: |:     |     |  |  |||: | : ||:  |:| : ||| | | |  |:
Db         151 WLRDGQKVESGVLTKPVEAETKGAGPATFSISSMLTITESDWLSQSLYTCRVDHRGIFFD 210

Qy         117 DS---SRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKK 173
                :   |  ||      |:  : | ||  | : : :| ||| || :  :: ::|     |
Db         211 KNVSMSSECSTTPSPGIQVFPIAPSFADTFLSKSARLICLVTDLTTYGSLNISWASHNGK 270

Qy         174 SIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKR 233
               :: :    :  | :|| |      | |:||  || : | | |   | |:  :|:|   :
Db         271 ALDTHMNITESHPNATFSAMGEASVCAEDWESGEQFTCTVTHADLPFPLKHTISK--SRE 328

Qy         234 LA---PEVYMLPPSPEE--TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQ 288
               :|   | ||:|||: |:     : |||||::|| |::: |||    :   :  : |: |
Db         329 VAKHPPAVYVLPPAREQLVLRESATVTCLVKGFSPADVFVQWQQRGQPLSSDKYVTSAPA 388

Qy         289 KDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALP---GSRTLEKSLHYSAG 340
               :      :| :| : |  :   | |  || | |||||     ||::||     |
Db         389 PEPQAPGLYFTHSTLTVTEEDWNSGETFTCVVGHEALPHMVTERTVDKSTEGEVG 443


RESULT 11
S37768
Ig mu chain C region - human
C;Species: Homo sapiens (man)
C;Date: 12-Feb-1998 #sequence_revision 12-Feb-1998 #text_change 23-Jul-1999
C;Accession: S37768
R;Harindranath, N.; Donadel, G.; Sigounas, G.; Notkins, A.L.
Mol. Immunol. 30, 111-112, 1993
```

```
A;Title: Comparison of complete nucleotide sequence of the human IgM heavy chain constan
A;Reference number: S37767; MUID:93109369; PMID:8417370
A;Accession: S37768
A;Molecule type: mRNA
A;Residues: 1-453 <HAR>
A;Cross-references: EMBL:X67301; NID:g38407; PIDN:CAA47714.1; PID:g38408
A;Experimental source: cell line Ab 63
C;Genetics:
A;Map position: 14q32
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: glycoprotein; heterotetramer; immunoglobulin
F;21-90/Domain: immunoglobulin homology <IMM1>
F;127-199/Domain: immunoglobulin homology <IMM2>
F;237-305/Domain: immunoglobulin homology <IMM3>
F;344-415/Domain: immunoglobulin homology <IMM4>
F;14/Disulfide bonds: interchain (to light chain) #status predicted
F;28-88,134-197,244-303,351-413/Disulfide bonds: #status predicted
F;46,209,272,279,440/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;214,452/Disulfide bonds: interchain (to heavy chain) #status predicted
F;291/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

  Query Match             26.3%;  Score 488.5;  DB 2;  Length 453;
  Best Local Similarity   31.2%;  Pred. No. 1.4e-28;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|    || | :|      ||     |:   : :|:|  :|||| :: |:|| :|::  :
Db        102 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 157

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                ||       |  |  |: : | : | :  |:| : :|| | | |  | :::|  |
Db        158 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 215

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
              |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        216 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 275

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        276 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 335

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :     :| :
Db        336 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 395

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        396 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 432


RESULT 12
S15590
Ig heavy chain - human
C;Species: Homo sapiens (man)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S15590
R;Neale, G.A.M.; Kitchingman, G.R.
Nucleic Acids Res. 19, 2427-2433, 1991
A;Title: mRNA transcripts initiating within the human immunoglobulin mu heavy chain enha
A;Reference number: S15590; MUID:91252286; PMID:1904154
A;Accession: S15590
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-474 <NEA>
A;Cross-references: EMBL:X58529
A;Note: the authors translated the codon CAA for residue 265 as Glu
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;344-415/Domain: immunoglobulin homology <IMM>

  Query Match             26.3%;  Score 488.5;  DB 2;  Length 474;
  Best Local Similarity   31.2%;  Pred. No. 1.5e-28;
  Matches  105;  Conservative   73;  Mismatches  134;  Indels   25;  Gaps   10;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|    || | :|      ||     |:   : :|:|  :|||| :: |:|| :|::  :
Db        102 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 157

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                ||       |  |  |: : | : | :  |:| : :|| | | |  | :::|  |
Db        158 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCV 215

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
              |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        216 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGQAVKTHTNISE 275

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        276 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 335

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :     :| :
Db        336 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 395

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        396 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 432


RESULT 13
MHRB
Ig mu chain C region, secreted form - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 23-Aug-1997
C;Accession: A02164
R;Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.
J. Immunol. 132, 490-495, 1984
A;Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-Ha2 a
A;Reference number: A02164; MUID:84088930; PMID:6418803
A;Contents: a2 allotype
A;Accession: A02164
A;Molecule type: mRNA
A;Residues: 1-458 <BER>
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F;21-92/Domain: immunoglobulin homology <IMM1>
F;130-202/Domain: immunoglobulin homology <IMM2>
F;242-310/Domain: immunoglobulin homology <IMM3>
F;349-420/Domain: immunoglobulin homology <IMM4>
F;14/Disulfide bonds: interchain (to light chain) #status predicted
F;28-90,137-200,249-308,356-418/Disulfide bonds: #status predicted
F;46,114,212,261,277,284,445/Binding site: carbohydrate (Asn) (covalent) #status predict
F;219,457/Disulfide bonds: interchain (to heavy chain) #status predicted
F;296/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

  Query Match             26.3%;  Score 488;  DB 1;  Length 458;
  Best Local Similarity   32.7%;  Pred. No. 1.5e-28;
  Matches  114;  Conservative   62;  Mismatches  149;  Indels   24;  Gaps    9;

Qy          5 HHHHTLSLPESGPV-TIIPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVT 59
              | :    |  | || : :|| | :|      ||     |       :|:|  :|||| :: |:
Db         94 HSNSNRDLRVSFPVDSELPPNVSVF---IPPRDSFSGSGTRKSRLICQATGFSPKQISVS 150

Qy         60 WLVDGQEAEN---LFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE 116
              || |||: |:     |    |  |  |||: | : ||:   |:| : ||| | | |  |:
Db        151 WLRDGQKVESGVLTKPVEAETKGAGPATFSISSMLTITESDWLSQSLYTCRVDHRGIFFD 210

Qy        117 DS---SRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKK 173
               :   |  ||      |:  : | ||  | : : : :|. ||| || :  :: ::|     |
Db        211 KNVSMSSECSTTPSPGIQVFPIAPSFADTFLSKSARLICLVTDLTTYGSLNISWASHNGK 270

Qy        174 SIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKR 233
              :: :    :  | :|| |       | |:||  || : | | |   | |:  :|:|   :
Db        271 ALDTHMNITESHPNATFSAMGEASVCAEDWESGEQFTCTVTHADLPFPLKHTISK--SRE 328
```

```
Qy        234 LA---PEVYMLPPSPEE--TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQ 288
              :|   | ||:|||: |:     : |||||::|| |::: |||    :   :  : |: |
Db        329 VAKHPPAVYVLPPAREQLVLRESATVTCLVKGFSPADVFVQWQQRGQPLSSDKYVTSAPA 388

Qy        289 KDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALP---GSRTLEKS 334
               :      :| :| : | :  |  |   || | |||||     ||::||
Db        389 PEPQAPGLYFTHSTLTVTEEDWNSGETFTCVVGHEALPHMVTERTVDKS 437


RESULT 14
MHHUBT
Ig mu heavy chain disease protein (Bot) - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
C;Accession: A02163
R;Barnikol-Watanabe, S.; Mihaesco, E.; Mihaesco, C.; Barnikol, H.U.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 365, 105-118, 1984
A;Title: The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid sequ
A;Reference number: A02163; MUID:84184186; PMID:6425189
A;Accession: A02163
A;Molecule type: protein
A;Residues: 1-391 <BAR>
C;Comment: This protein has no V region homology or CH1 region.
C;Genetics:
A;Gene: GDB:IGHM
A;Cross-references: GDB:120086; OMIM:147020
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin; transmembrane pro
F;1-41/Domain: pre-C <VAR>
F;43-391/Domain: Ig mu chain C region, secreted form <IGM>
F;65-137/Domain: immunoglobulin homology <IMM1>
F;175-243/Domain: immunoglobulin homology <IMM2>
F;282-353/Domain: immunoglobulin homology <IMM3>
F;147,210,217,378/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             26.1%;  Score 484.5;  DB 1;  Length 391;
  Best Local Similarity   31.2%;  Pred. No. 2.2e-28;
  Matches  105;  Conservative   72;  Mismatches  137;  Indels   23;  Gaps   10;

Qy         15 SGPVTI-IPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              | ||   :|| | :|      ||     |:   : :|:|  :|||| :: |:|| :|::   :
Db         40 SQPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIEVSWLREGKQVGS 95

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||          |  |  |: : | : | :  |:  : :|| | | |  | :::|  |
Db         96 --GVTTDEVEAEAKESGPTTYKVTSTLTIKESDWLGQSMFTCRVDHRGLTFQQNASSMCG 153

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
              |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        154 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQDGEAVKTHTNISE 213

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
              | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        214 SHPNATFSAVGEASICEDDWDSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 273

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :     :| :
Db        274 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 333

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        334 SILTVSEEEWNTGETYTCVVAHEALPNRVTERTVDKS 370


RESULT 15
MHHU
Ig mu chain C region, secreted splice form - human
C;Species: Homo sapiens (man)
C;Date: 29-Jul-1981 #sequence_revision 23-Aug-1997 #text_change 22-Jun-1999
C;Accession: S09357; S16656; S37767; A26243; A26244; I37749; I37750; A02162; B02162
R;Dorai, H.; Gillies, S.D.
Nucleic Acids Res. 17, 6412, 1989
A;Title: The complete nucleotide sequence of a human immunoglobulin genomic C-mu gene.
A;Reference number: S09357; MUID:89366690; PMID:2505237
A;Accession: S09357
A;Molecule type: DNA
A;Residues: 1-452 <DOR1>
A;Cross-references: EMBL:X14940
A;Note: the authors translated the codon AAT for residue 16 as Met
R;Dorai, H.
submitted to the EMBL Data Library, April 1989
A;Reference number: S16656
A;Accession: S16656
A;Molecule type: DNA
A;Residues: 1-39,'L',41-452 <DOR2>
A;Cross-references: EMBL:X14940
R;Harindranath, N.; Donadel, G.; Sigounas, G.; Notkins, A.L.
Mol. Immunol. 30, 111-112, 1993
A;Title: Comparison of complete nucleotide sequence of the human IgM heavy chain constan
A;Reference number: S37767; MUID:93109369; PMID:8417370
A;Accession: S37767
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-190,'S',192-414,'V',415-452 <HAR>
A;Cross-references: EMBL:X67292; NID:g38405; PIDN:CAA47708.1; PID:g38406
R;Rabbitts, T.H.; Forster, A.; Milstein, C.P.
Nucleic Acids Res. 9, 4509-4524, 1981
A;Title: Human immunoglobulin heavy chain genes: evolutionary comparisons of C-mu, C-del
A;Reference number: A26243; MUID:82059479; PMID:6795593
A;Accession: A26243
A;Molecule type: DNA
A;Residues: 1-17;105-186;200-259;296-322;339-416,'D',418-452 <RAB>
A;Cross-references: GB:K01310; NID:g184715
R;Dolby, T.W.; Devuono, J.; Croce, C.M.
Proc. Natl. Acad. Sci. U.S.A. 77, 6027-6031, 1980
A;Title: Cloning and partial nucleotide sequence of human immunoglobulin mu chain cDNA f
A;Reference number: A26244; MUID:81077306; PMID:6777778
A;Accession: A26244
A;Molecule type: mRNA
A;Residues: 298-386;436-452 <DOL>
A;Cross-references: GB:J00257; NID:g185053; PIDN:AAA53508.1; PID:g185056; GB:J00258; NID
R;Takahashi, N.; Nakai, S.; Honjo, T.
Nucleic Acids Res. 8, 5983-5991, 1980
A;Title: Cloning of human immunoglobulin mu gene and comparison with mouse mu gene.
A;Reference number: I37748; MUID:81124312; PMID:6450943
A;Accession: I37749
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 347-370 <TAK1>
A;Cross-references: EMBL:V00562; NID:g33448; PIDN:CAA23825.1; PID:g929649
A;Accession: I37750
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 433-452 <TAK2>
A;Cross-references: EMBL:V00563; NID:g33454; PIDN:CAA23826.1; PID:g825684
R;Mihaesco, E.; Barnikol-Watanabe, S.; Barnikol, H.U.; Mihaesco, C.; Hilschmann, N.
Eur. J. Biochem. 111, 275-286, 1980
A;Title: The primary structure of the constant part of mu-chain-disease protein BOT.
A;Reference number: A02162; MUID:81066716; PMID:6777162
A;Contents: Mu-chain-disease protein Bot and revisions to sequence of Gal
A;Accession: A02162
A;Molecule type: protein
A;Residues: 1-17,'BPS',22,'T',24-82,'N',84-90,'Z',92-93,'B',95,'B',97-144,'E',146-162,'E
A;Accession: B02162
A;Molecule type: protein
A;Residues: 100-144,'E',146-162,'E',164,'E',166-214,'G',216-262,'D',264-295,'D',297-414,
A;Note: all four combinations of the 191-Ser/Gly and 215-Val/Gly polymorphisms have been
R;Watanabe, S.; Barnikol, H.U.; Horn, J.; Bertram, J.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 354, 1505-1509, 1973
A;Title: The primary structure of a monoclonal IgM-immunoglobulin (macroglobulin Gal), I
.
A;Reference number: A02064; MUID:75059123; PMID:4803843
A;Contents: annotation; Waldenstrom's macroglobulin Gal
A;Note: this sequence has been revised in reference A02162
```

```
R;Putnam, F.W.; Florent, G.; Paul, C.; Shinoda, T.; Shimizu, A.
Science 182, 287-291, 1973
A;Title: Complete amino acid sequence of the mu heavy chain of a human IgM immunoglobuli
A;Reference number: A02088; MUID:74005511; PMID:4742735
A;Contents: annotation; Waldenstrom's macroglobulin Ou, sequence, disulfide bonds, and c
A;Note: this sequence differs from that shown at a number of positions; this sequence ha
C;Comment: During differentiation, B lymphocytes switch from expression of membrane-boun
s.
C;Genetics:
A;Gene: GDB:IGHM
A;Cross-references: GDB:120086; OMIM:147020
A;Map position: 14q32.33-14q32.33
A;Introns: 1/1; 105/1; 217/1; 323/1; 433/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. The IgM subunits associate into disulfide linked pentamers.
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: alternative splicing; glycoprotein; heterotetramer; immunoglobulin
F;21-90/Domain: immunoglobulin homology <IMM1>
F;127-199/Domain: immunoglobulin homology <IMM2>
F;237-305/Domain: immunoglobulin homology <IMM3>
F;344-415/Domain: immunoglobulin homology <IMM4>
F;432-452/Domain: carboxyl-terminal <CTS>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;28-88,134-197,244-303,351-413/Disulfide bonds: #status experimental
F;46,209,272,279,439/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;214,451/Disulfide bonds: interchain (to heavy chain) #status experimental
F;291/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental

  Query Match             25.7%;  Score 477;  DB 1;  Length 452;
  Best Local Similarity   30.9%;  Pred. No. 9.7e-28;
  Matches  104;  Conservative   73;  Mismatches  134;  Indels   26;  Gaps   11;

Qy         17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAEN 69
              |: :|   || | :|     ||    |:   : :|:|  :|||| :: |:|| :|::  :
Db        102 PLPVIAELPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS 157

Qy         70 LFPYTT-----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCS 123
                 ||        |  |  |: : | : | :  |:  : :|| | | |  | :::|  |
Db        158 --GVTTDQVQAEAKESGPTTYKVTSTLTIKESDWLGQSMFTCRVDHRGLTFQQNASSMCV 215

Qy        124 DDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRST 183
               |:   :  : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |
Db        216 PDQDTAIRVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISE 275

Qy        184 KHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLP 242
               | :|| |      :   ||  || : | | |   | |: ::|::  |  |  |:||:||
Db        276 SHPNATFSAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLP 335

Qy        243 PSPEETG--TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLY 300
              |: |:     : |:|||: || |::: |||:   :      : |: |  :      :| :
Db        336 PAREQLNLRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAH 395

Qy        301 SRMLVNKSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
              | : |::  |  |   || | |||||     ||::||
Db        396 SILTVSEEEWNTGETYTC-VAHEALPNRVTERTVDKS 431


Search completed: July  9, 2003, 13:39:44
Job time : 21 secs
```

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 15 Seconds
(without alignments)
668.881 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHHTLSLPESGPVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	950.5	51.2	426	1	US-08-336-583-2	Sequence 2, Appli
2	950.5	51.2	426	5	PCT-US95-13795-2	Sequence 2, Appli
3	862.5	46.4	561	3	US-09-192-545-2	Sequence 2, Appli
4	507	27.3	331	2	US-08-646-981-17	Sequence 17, Appl
5	493.5	26.6	334	2	US-08-646-981-16	Sequence 16, Appl
6	488	26.3	504	1	US-07-932-915-2	Sequence 2, Appli
7	488	26.3	504	5	PCT-US91-05826-2	Sequence 2, Appli
8	475.5	25.6	333	1	US-08-436-463-6	Sequence 6, Appli
9	475.5	25.6	333	1	US-08-024-253-6	Sequence 6, Appli
10	475	25.6	450	2	US-08-788-800-12	Sequence 12, Appl
11	475	25.6	469	2	US-07-934-373C-23	Sequence 23, Appl
12	475	25.6	469	3	US-08-437-642B-23	Sequence 23, Appl
13	475	25.6	469	4	US-08-146-206C-23	Sequence 23, Appl
14	469.5	25.3	530	3	US-08-477-460B-4	Sequence 4, Appli
15	469.5	25.3	530	3	US-08-379-516-4	Sequence 4, Appli
16	469.5	25.3	530	4	US-09-329-916-4	Sequence 4, Appli
17	469.5	25.3	530	4	US-08-485-372A-4	Sequence 4, Appli
18	469.5	25.3	530	4	US-09-409-006A-4	Sequence 4, Appli
19	469.5	25.3	530	4	US-08-484-681-4	Sequence 4, Appli
20	469.5	25.3	530	5	PCT-US93-07422-4	Sequence 4, Appli
21	457	24.6	326	2	US-08-656-586-9	Sequence 9, Appli
22	454	24.4	552	5	PCT-US93-07832-23	Sequence 23, Appl
23	451.5	24.3	327	2	US-08-761-277A-47	Sequence 47, Appl
24	451.5	24.3	443	5	PCT-US96-13152-4	Sequence 4, Appli
25	451.5	24.3	467	1	US-08-704-744-81	Sequence 81, Appl
26	451.5	24.3	467	4	US-08-523-894-8	Sequence 8, Appli
27	451.5	24.3	467	4	US-08-523-894-10	Sequence 10, Appl
28	451.5	24.3	467	4	US-08-523-894-12	Sequence 12, Appl
29	447	24.1	451	4	US-09-247-352-3	Sequence 3, Appli
30	447	24.1	451	4	US-09-466-635-3	Sequence 3, Appli
31	447	24.1	472	4	US-09-301-593-30	Sequence 30, Appl
32	447	24.1	472	4	US-09-301-593-43	Sequence 43, Appl
33	446.5	24.0	451	2	US-08-887-352B-14	Sequence 14, Appl
34	446.5	24.0	451	2	US-08-887-352B-16	Sequence 16, Appl
35	446.5	24.0	451	2	US-08-887-352B-18	Sequence 18, Appl
36	446.5	24.0	451	3	US-08-466-151-65	Sequence 65, Appl
37	446.5	24.0	451	4	US-09-109-207C-14	Sequence 14, Appl
38	446.5	24.0	451	4	US-09-109-207C-16	Sequence 16, Appl
39	446.5	24.0	451	4	US-09-109-207C-18	Sequence 18, Appl
40	446.5	24.0	451	4	US-09-282-505-2	Sequence 2, Appli
41	446.5	24.0	451	4	US-09-054-255-2	Sequence 2, Appli
42	446.5	24.0	451	4	US-09-296-005-14	Sequence 14, Appl
43	446.5	24.0	451	4	US-09-296-005-16	Sequence 16, Appl
44	446.5	24.0	451	4	US-09-296-005-18	Sequence 18, Appl
45	445.5	24.0	453	3	US-08-466-151-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-336-583-2
; Sequence 2, Application US/08336583
; Patent No. 5629415
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,583
; FILING DATE: 09-NOV-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-583-2

Query Match 51.2%; Score 950.5; DB 1; Length 426;
Best Local Similarity 53.8%; Pred. No. 2.2e-86;
Matches 175; Conservative 52; Mismatches 95; Indels 3; Gaps 2;

```
Qy          18 VTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLPPYTTRP 77
               :  |||||||||||:| || |:|||||||:||: |  : | ||||||:| |:||||
Db         102 LNFIPPTVKLFHSSCNPVGDTHTTIQLLCLISGYVPGDMEVIWLVDGQKATNIFPYTAPG 161

Qy          78 KREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIPP 137
```

```
Db        162 TKEGNVT-STHSELNITQGEWVSQKTYTCQVTYQGFTFKDEARKCSESDPRGVTSYLSPP 220

Qy        138 SPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITSILP 197

Db        221 SPLDLYVHKAPKITCLVVDLATMEGMNLTWYRESKEPVNPGPLNKKDHFNGTITVTSTLP 280

Qy        198 VDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR--TVT 255

Db        281 VNTNDWIEGETYYCRVTHPHLPKDIVRSIAKAPGKRAPPDVYLFLPPEEEQGTKDRVTLT 340

Qy        256 CLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNL 315

Db        341 CLIQNFFPADISVQWLRNDSPIQTDQYTTTGPHKVSGSRPAFFIFSRLEVSRVDWEQKNK 400

Qy        316 VTCRVVHEALPGSRTLEKSLHYSAG 340

Db        401 FTCQVVHEALSGSRILQKWVSKTPG 425


RESULT 2
PCT-US95-13795-2
; Sequence 2, Application PC/TUS9513795
;  GENERAL INFORMATION:
;    APPLICANT:  HOLLIS, GREGORY F.
;    APPLICANT:  PATEL, MAYUR D.
;    TITLE OF INVENTION:  DNA ENCODING CANINE IMMUNOGLOBULINS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHRISTINE E. CARTY
;      STREET:  126 E. LINCOLN AVENUE; P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07065-0907
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13795
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CARTY, CHRISTINE E.
;      REGISTRATION NUMBER:  36,099
;      REFERENCE/DOCKET NUMBER:  19211Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-6734
;      TELEFAX:  (908) 594-4720
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  426 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-13795-2

  Query Match             51.2%;  Score 950.5;  DB 5;  Length 426;
  Best Local Similarity   53.8%;  Pred. No. 2.2e-86;
  Matches  175;  Conservative   52;  Mismatches   95;  Indels    3;  Gaps    2;

Qy         18 VTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRP 77

Db        102 LNFIPPTVKLFHSSCNPVGDTHTTIQLLCLISGYVPGDMEVIWLVDGQKATNIFPYTAPG 161

Qy         78 KREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVITYLIPP 137

Db        162 TKEGNVT-STHSELNITQGEWVSQKTYTCQVTYQGFTFKDEARKCSESDPRGVTSYLSPP 220

Qy        138 SPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATTSITSILP 197

Db        221 SPLDLYVHKAPKITCLVVDLATMEGMNLTWYRESKEPVNPGPLNKKDHFNGTITVTSTLP 280

Qy        198 VDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEETGTTR--TVT 255

Db        281 VNTNDWIEGETYYCRVTHPHLPKDIVRSIAKAPGKRAPPDVYLFLPPEEEQGTKDRVTLT 340

Qy        256 CLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNL 315

Db        341 CLIQNFFPADISVQWLRNDSPIQTDQYTTTGPHKVSGSRPAFFIFSRLEVSRVDWEQKNK 400

Qy        316 VTCRVVHEALPGSRTLEKSLHYSAG 340

Db        401 FTCQVVHEALSGSRILQKWVSKTPG 425


RESULT 3
US-09-192-545-2
; Sequence 2, Application US/09192545
; Patent No. 6118044
; GENERAL INFORMATION:
;  APPLICANT: Karasuyama, Hajime
;  APPLICANT: Yonekawa, Hiromichi
;  APPLICANT: Taya, Choji
;  APPLICANT: Matsuoka, Kunie
;  TITLE OF INVENTION: Transgenic Animal Allergy Models and Methods for Their Use
;  FILE REFERENCE: 799P79570
;  CURRENT APPLICATION NUMBER: US/09/192,545
;  CURRENT FILING DATE: 1998-11-13
;  EARLIER APPLICATION NUMBER: JP HEI 9-313989
;  EARLIER FILING DATE: 1997-11-14
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 561
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE: Description of Artificial Sequence: Designed heavy
US-09-192-545-2

  Query Match             46.4%;  Score 862.5;  DB 3;  Length 561;
  Best Local Similarity   51.8%;  Pred. No. 1.9e-77;
  Matches  183;  Conservative   46;  Mismatches  103;  Indels   21;  Gaps    8;

Qy          2 FHHHHHHTLSLPESGPVTIIPPTVKLFHS----------SCDPRGDAHSTIQLLCLVSGF 51

Db        211 FTCHVTHPPSFNESRTILVRPVT----HSLSPPWSYSIHRCDPNA-FHSTIQLYCFIYGH 265

Qy         52 SPAKVHVTWLVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHN 111

Db        266 ILNDVSVSWLMDDREITDTLAQTVLIKEE-GKLASTCSKLNITEQQWMSESTFTCRVTSQ 324

Qy        112 GSIFEDSSRRCSDDEPRGVITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRER 171

Db        325 GVDYLAHTRRCPDHEPRGAITYLIPPSPLDLYQNGAPKLTCLVVDLESEKNVNVTWNQEK 384

Qy        172 KKSIGSASQRSTKHH-HATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLP 230

Db        385 KTSV-SASQWYTKHHNNATTSITSILPVVAKDWIEGYGYQCVVDRPDFPKPIVRSIT-LP 442

Qy        231 --GKRLAPEVYMLPPSPEETGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQ 288

Db        443 QVSQRSAPEVYVFPPPEEESEDKRTLTCLIQNFFPEDISVQWLGDGKLISNSQHSTTTPL 502

Qy        289 KDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAGN 341

Db        503 KSNGSNQGFFIFSRLEVAKTLWTQRKQFTCQVIHEALQKPRKLEKTISTSLGN 555


RESULT 4
US-08-646-981-17
; Sequence 17, Application US/08646981
; Patent No. 5852183
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  MAEDA, HIROAKI
;    APPLICANT:  EDA, YASUYUKI
;    APPLICANT:  KIMACHI, KAZUHIKO
;    APPLICANT:  ONO, YOICHI
;    APPLICANT:  TOKIYOSHI, SACHIO
;    TITLE OF INVENTION:  DOG-MOUSE HETEROHYBRIDOMA AND GENE
;    TITLE OF INVENTION:  FRAGMENT CODING FOR CONSTANT REGION OF CANINE
;    TITLE OF INVENTION:  IMMUNOGLOBULINS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/646,981
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEINER, MARC S
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  1488-106
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  331 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-646-981-17

  Query Match             27.3%;  Score 507;  DB 2;  Length 331;
  Best Local Similarity   35.2%;  Pred. No. 2.3e-42;
  Matches  118;  Conservative   69;  Mismatches  122;  Indels   26;  Gaps   11;

Qy         23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTTRPKREGG 82
              |:|     ||     : ||: | |||||: |  | |:|   |      :  : :  | |
Db          5 PSVFPLAPSCG--STSGSTVTLACLVSGYFPEPVTVSW-NSGSLTSGVHTFPSVLKSSG- 60

Qy         83 QTFSLQSEVNITQGQWMSSNTYTCHVKHNGS--------IFEDSSRRCSDDEP-----RG 129
                :|| | | :   : : | |:||:| | :        : ::|: :|    |       |
Db         61 -LYSLSSMVTVPSSR-LPSETFTCNVVHPATNTKVDKPGVPKESTCKCISPCPVPESLGG 118

Qy        130 VITYLIPPSPLDLYE-NGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHH 187
               :: || | |:     ||::||:|||| |:  : ::|  : |:   : :|   :  :
Db        119 PSVFIFPPFPKDILRITRTPEVTCVVLDLGREDPEVQISWFVDGKEVHTAKTQPREQQFN 178

Qy        188 ATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE 247
               :|  : |:||:: :||: |: ::|||:|    | || |:|:|  |:   | ||:|||||:|
Db        179 STYRVVSVLPIEHQDWLTGKEFKCRVNHIGLPSPIERTISKARGQAHQPGVYVLPPSPKE 238

Qy        248 --TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLV 305
                 :  | |:|||: |:| || |:|  | : :    : || || |   | |:||||:: |
Db        239 LSSSDTVTLTCLIKDFFPPEIDVEWQSNGQPEPESKYHTTAPQLDE--DGSYFLYSKLSV 296

Qy        306 NKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              :|| ||:|:  || |:|||    | : || :| |
Db        297 DKSRWEQGDPFTCAVMHEALQNHYT-DLSLSHSPG 330


RESULT 5
US-08-646-981-16
; Sequence 16, Application US/08646981
; Patent No. 5852183
;  GENERAL INFORMATION:
;    APPLICANT:  MAEDA, HIROAKI
;    APPLICANT:  EDA, YASUYUKI
;    APPLICANT:  KIMACHI, KAZUHIKO
;    APPLICANT:  ONO, YOICHI
;    APPLICANT:  TOKIYOSHI, SACHIO
;    TITLE OF INVENTION:  DOG-MOUSE HETEROHYBRIDOMA AND GENE
;    TITLE OF INVENTION:  FRAGMENT CODING FOR CONSTANT REGION OF CANINE
;    TITLE OF INVENTION:  IMMUNOGLOBULINS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/646,981
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEINER, MARC S
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  1488-106
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  334 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-646-981-16

  Query Match             26.6%;  Score 493.5;  DB 2;  Length 334;
  Best Local Similarity   33.9%;  Pred. No. 5.3e-41;
  Matches  116;  Conservative   68;  Mismatches  121;  Indels   37;  Gaps   11;

Qy         23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPK 78
              |:|     ||     : ||: | |||||: |  | |:|    |  |         :|
Db          5 PSVFPLDPSCG--STSGSTVALACLVSGYFPEPVTVSWNSGSLTSG-------VHTFPSD 55

Qy         79 REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGS---IFEDSSRRCSDDEPR------- 128
                :    :|| | | :   :| || |:||:| |  |   : :   :| :   ||
Db         56 LQSSGLYSLSSMVTVPSSRW-SSETFTCNVAHPASKTKVDKPVPKRENGRVPRPPDCPKC 114

Qy        129 -------GVITYLIPPSPLD-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSAS 179
                    |   :: || | | |     ||::||:|:||  |:  : ::|  : |:   : :
Db        115 PAPEMLGGPSVFIFPPKPKDTLLIARTPEVTCVVVDLGPEDPEVQISWFVDGKQMQTAKT 174

Qy        180 QRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
              |   :  : |  : |:||:  :||::|: : |:|::   | || |:|:|  |:   | ||
Db        175 QPREEQFNGTYRVVSVLPIGHQDWLKGKQFTCKVNNKALPSPIERTISKARGQAHQPSVY 234

Qy        240 MLPPSPEE-TGTTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFF 298
              :|||| || :  | ::|||||: |:| :| |:|  | :::    : || || |   | |:|
Db        235 VLPPSREELSKNTVSLTCLIKDFFPPDIDVEWQSNGQQEPESKYRTTPPQLDE--DGSYF 292

Qy        299 LYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              |||:: |:|| |::|:   | |:||||    | :||| :| |
Db        293 LYSKLSVDKSRWQRGDTFICAVMHEALHNHYT-QKSLSHSPG 333


RESULT 6
US-07-932-915-2
; Sequence 2, Application US/07932915
; Patent No. 5672486
;  GENERAL INFORMATION:
;    APPLICANT:  Soulillou, Jean-Paul
```

```
;    TITLE OF INVENTION:  Protein Polyligands Joined To A Stable Protein
;    TITLE OF INVENTION:  Core
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square, Suite 400
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/932,915
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/646,875
;      FILING DATE:  28-JAN-1991
;      APPLICATION NUMBER:  US 07/575,394
;      FILING DATE:  23-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland Ph.D., Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  ATLA-001/01US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-494-7622
;      TELEFAX:  415-857-0663
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-932-915-2

  Query Match             26.3%;  Score 488;  DB 1;  Length 504;
  Best Local Similarity   31.2%;  Pred. No. 3.5e-40;
  Matches  103;  Conservative   72;  Mismatches  133;  Indels   22;  Gaps    9;

Qy         21 IPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTT- 75
              :|| | :|     ||    |:  : :|:|  :|||| :: |:|| :|::  :    ||
Db        160 LPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS--GVTTD 213

Qy         76 ----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCSDDEPRGV 130
                  |  |  |: : | : | :  |:| : :|| | | |  | :::|  |  |:  :
Db        214 QVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCVPDQDTAI 273

Qy        131 ITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATT 190
               : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |  | :||
Db        274 RVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISESHPNATF 333

Qy        191 SITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLPPSPEETG 249
              |      :   ||  || : | | |   | |: ::|::  |  |  |:||:|||: |:
Db        334 SAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLPPAREQLN 393

Qy        250 --TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNK 307
                : |:|||: || |::: |||:   :      : |: |  :      :| :| : |::
Db        394 LRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAHSILTVSE 453

Qy        308 SIWEKGNLVTCRVVHEALPG---SRTLEKS 334
                |  |   || | |||||     ||::||
Db        454 EEWNTGETYTCVVAHEALPNRVTERTVDKS 483


RESULT 7
PCT-US91-05826-2
; Sequence 2, Application PC/TUS9105826
;  GENERAL INFORMATION:
;    APPLICANT:  Soulillou, Jean-Paul
```

```
;    TITLE OF INVENTION:  Protein Polyligands Joined To A Stable Protein
;    TITLE OF INVENTION:  Core
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bertram I. Rowland, Ph.D.
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/05826
;      FILING DATE:  19910822
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/575,394
;      FILING DATE:  29-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland Ph.D., Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  FP55352-1/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-494-8771
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-05826-2

  Query Match             26.3%;  Score 488;  DB 5;  Length 504;
  Best Local Similarity   31.2%;  Pred. No. 3.5e-40;
  Matches  103;  Conservative   72;  Mismatches  133;  Indels   22;  Gaps    9;

Qy         21 IPPTVKLFHSSCDPR----GDAHSTIQLLCLVSGFSPAKVHVTWLVDGQEAENLFPYTT- 75
              :|| | :|     ||    |:  : :|:|  :|||| :: |:|| :|::  :    ||
Db        160 LPPKVSVF---VPPRDGFFGNPRKS-KLICQATGFSPRQIQVSWLREGKQVGS--GVTTD 213

Qy         76 ----RPKREGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIF-EDSSRRCSDDEPRGV 130
                  |  |  |: : | : | :  |:| : :|| | | |  | :::|  |  |:  :
Db        214 QVQAEAKESGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFQQNASSMCVPDQDTAI 273

Qy        131 ITYLIPPSPLDLYENGTPKLTCLVLDLESEENITVTWVRERKKSIGSASQRSTKHHHATT 190
               : ||||   ::   : |||||| || : :::|::| |:  ::: : :  |  | :||
Db        274 RVFAIPPSFASIFLTKSTKLTCLVTDLTTYDSVTISWTRQNGEAVKTHTNISESHPNATF 333

Qy        191 SITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRL-APEVYMLPPSPEETG 249
              |      :   ||  || : | | |   | |: ::|::  |  |  |:||:|||: |:
Db        334 SAVGEASICEDDWNSGERFTCTVTHTDLPSPLKQTISRPKGVALHRPDVYLLPPAREQLN 393

Qy        250 --TTRTVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNK 307
                : |:|||: || |::: |||:   :      : |: |  :      :| :| : |::
Db        394 LRESATITCLVTGFSPADVFVQWMQRGQPLSPEKYVTSAPMPEPQAPGRYFAHSILTVSE 453

Qy        308 SIWEKGNLVTCRVVHEALPG---SRTLEKS 334
                |  |   || | |||||     ||::||
Db        454 EEWNTGETYTCVVAHEALPNRVTERTVDKS 483


RESULT 8
US-08-436-463-6
; Sequence 6, Application US/08436463
; Patent No. 5760185
;  GENERAL INFORMATION:
;    APPLICANT:  KIMACHI, Kazuhiko
;    APPLICANT:  MAEDA, Hiroaki
```

```
;    APPLICANT:  NISHIYAMA, Kiyoto
;    APPLICANT:  TOKIYOSHI, Sachio
;    TITLE OF INVENTION:  ANTI-FELINE HERPES VIRUS-1 RECOMBINANT
;    TITLE OF INVENTION:  ANTIBODY AND GENE FRAGMENT CODING FOR SAID ANTIBODY
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  419 Seventh Street, N.W., Suite 400
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,463
;      FILING DATE:  26-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 341255/1992
;      FILING DATE:  28-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YUN, Allen C.
;      REGISTRATION NUMBER:  37,971
;      REFERENCE/DOCKET NUMBER:  KIMACHI=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  333 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-436-463-6

  Query Match             25.6%;  Score 475.5;  DB 1;  Length 333;
  Best Local Similarity   32.7%;  Pred. No. 3.3e-39;
  Matches  112;  Conservative   72;  Mismatches  121;  Indels   37;  Gaps   12;

Qy         23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPK 78
              |:|     ||      |  |: | ||| |: |  | |:|     | |    : ||
Db          4 PSVFPLAPSCGTTSGA--TVALACLVLGYFPEPVTVSWNSGALTSG---VHTFPAVL--- 55

Qy         79 REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE-DSSRRCSDDEP---------- 127
               :   :|| | | :  :|: |:|:||:| |  |  : | : | :|  |
Db         56 -QASGLYSLSSMVTVPSSRWL-SDTFTCNVAHPPSNTKVDKTVRKTDHPPGPKPCDCPKC 113

Qy        128 ------RGVITYLIPPSPLD-LYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSAS 179
                   |   :: || | | |  : ||::||||:||   : :: :||  :  :   : :
Db        114 PPPEMLGGPSIFIFPPKPKDTLSISRTPEVTCLVVDLGPDDSDVQITWFVDNTQVYTAKT 173

Qy        180 QRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                  :  ::|  : |:||:  :||::|: ::|:|:    | || |:|:|  |:   |:||
Db        174 SPREEQFNSTYRVVSVLPILHQDWLKGKEFKCKVNSKSLPSPIERTISKAKGQPHEPQVY 233

Qy        240 MLPPSPEETGTTR-TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFF 298
               :|||: ||    : :|||||: |:| :|:|:|    : :   :: || || |  :| ::|
Db        234 VLPPAQEELSRNKVSVTCLIKSFHPPDIAVEWEITGQPEPENNYRTTPPQLD--SDGTYF 291

Qy        299 LYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              :||:: |::| |::||  || | ||||    | :|||  | |
Db        292 VYSKLSVDRSHWQRGNTYTCSVSHEALHSHHT-QKSLTQSPG 332


RESULT 9
US-08-024-253-6
; Sequence 6, Application US/08024253
; Patent No. 5785968
;  GENERAL INFORMATION:
;    APPLICANT:  KIMACHI, Kazuhiko
;    APPLICANT:  MAEDA, Hiroaki
;    APPLICANT:  NISHIYAMA, Kiyoto
;    APPLICANT:  TOKIYOSHI, Sachio
;    APPLICANT:  TOHYA, Yukinobu
;    APPLICANT:  MIKAMI, Takeshi
;    TITLE OF INVENTION:  ANTI-FELINE CALCIVIRUS RECOMBINANT
;    TITLE OF INVENTION:  ANTIBODY AND GENE FRAGMENT ENCODING THE SAME
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WEGNER, CANTOR, MUELLER & PLAYER
;      STREET:  1233 20th Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20036-8218
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/024,253
;      FILING DATE:  19930301
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 79189/1992
;      FILING DATE:  28-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CANTOR, Herbert I.
;      REGISTRATION NUMBER:  24,392
;      REFERENCE/DOCKET NUMBER:  P-500-23744
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-0400
;      TELEFAX:  (202) 835-0605
;      TELEX:  440706 WEGBR
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  333 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-024-253-6

  Query Match             25.6%;  Score 475.5;  DB 1;  Length 333;
  Best Local Similarity   32.7%;  Pred. No. 3.3e-39;
  Matches  112;  Conservative   72;  Mismatches  121;  Indels   37;  Gaps   12;

Qy         23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPK 78
              |:|     ||      |  |: | ||| |: |  | |:|     | |    : ||
Db          4 PSVFPLAPSCGTTSGA--TVALACLVLGYFPEPVTVSWNSGALTSG---VHTFPAVL--- 55

Qy         79 REGGQTFSLQSEVNITQGQWMSSNTYTCHVKHNGSIFE-DSSRRCSDDEP---------- 127
               :    :|| | | :   :|: |:|:||:| |  |  : | : | :|  |
Db         56 -QASGLYSLSSMVTVPSSRWL-SDTFTCNVAHPPSNTKVDKTVRKTDHPPGPKPCDCPKC 113

Qy        128 ------RGVITYLIPPSPLD-LYENGTPKLTCLVLDL-ESEENITVTWVRERKKSIGSAS 179
                   |   :: || | | |  : ||::||||:||   : :: :||  :  :   : :
Db        114 PPPEMLGGPSIFIFPPKPKDTLSISRTPEVTCLVVDLGPDDSDVQITWFVDNTQVYTAKT 173

Qy        180 QRSTKHHHATTSITSILPVDAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVY 239
                  :  ::|  : |:||:  :||::|: ::|:|:    | || |:|:|  |:   |:||
Db        174 SPREEQFNSTYRVVSVLPILHQDWLKGKEFKCKVNSKSLPSPIERTISKAKGQPHEPQVY 233

Qy        240 MLPPSPEETGTTR-TVTCLIRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFF 298
               :|||: ||    : :|||||: |:| :|:|:|    : :   :: || || |  :| ::|
Db        234 VLPPAQEELSRNKVSVTCLIKSFHPPDIAVEWEITGQPEPENNYRTTPPQLD--SDGTYF 291

Qy        299 LYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
              :||:: |::| |::||  || | ||||    | :|||  | |
Db        292 VYSKLSVDRSHWQRGNTYTCSVSHEALHSHHT-QKSLTQSPG 332
```

```
RESULT 10
US-08-788-800-12
; Sequence 12, Application US/08788800
; Patent No. 5914112
;  GENERAL INFORMATION:
;    APPLICANT:  Bednar, Martin M.
;    APPLICANT:  Thomas, G. Roger
;    APPLICANT:  Gross, Cordell E.
;    TITLE OF INVENTION:  ANTI-CD18 ANTIBODIES IN STROKE
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/788,800
;      FILING DATE:  22-Jan-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  40,378
;      REFERENCE/DOCKET NUMBER:  P0987r1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  450 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-788-800-12

  Query Match             25.6%;  Score 475;  DB 2;  Length 450;
  Best Local Similarity   34.7%;  Pred. No. 5.8e-39;
  Matches  112;  Conservative   66;  Mismatches  115;  Indels   30;  Gaps   12;

Qy         35 RGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSE 90
              |  : ||  | |||  : |  | |:|    |  |    : ||       :    :|| |
Db        140 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSV 192

Qy         91 VNITQGQWMSSNTYTCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPL 140
              | :|   :  : ||||:| |   |  : :   |:|  : |       |   :| || |
Db        193 VTVTSSNF-GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPK 251

Qy        141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPV 198
              | |  : ||::||:|:|:  |:  :   |  :  :   : ::   :  ::|  : |:| |
Db        252 DTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGMEVHNAKTKPREEQFNSTFRVVSVLTV 311

Qy        199 DAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCL 257
                :||: |: |:|:| :   | || ::|:|  |:   |:|| |||| || |    ::|||
Db        312 VHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCL 371

Qy        258 IRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVT 317
              ::|||||:|:|:|  | : ::  :: || |  |  :| ||||||:: |:|| |::||: :
Db        372 VKGFYPSDIAVEWESNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFS 427

Qy        318 CRVVHEALPGSRTLEKSLHYSAG 340
              | |:||||    | :|||  | |
Db        428 CSVMHEALHNHYT-QKSLSLSPG 449


RESULT 11
US-07-934-373C-23
; Sequence 23, Application US/07934373C
; Patent No. 5821337
;  GENERAL INFORMATION:
;    APPLICANT:  Paul J. Carter
;    APPLICANT:  Leonard G. Presta
;    TITLE OF INVENTION:  Immunoglobulin Variants
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/934,373C
;      FILING DATE:  21-Aug-1992
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/05126
;      FILING DATE:  15-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/715272
;      FILING DATE:  14-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  40,378
;      REFERENCE/DOCKET NUMBER:  P0709P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-1994
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  469 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-07-934-373C-23

  Query Match             25.6%;  Score 475;  DB 2;  Length 469;
  Best Local Similarity   34.7%;  Pred. No. 6.2e-39;
  Matches  112;  Conservative   66;  Mismatches  115;  Indels   30;  Gaps   12;

Qy         35 RGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSE 90
              |  : ||  | |||  : |  | |:|    |  |    : ||       :    :|| |
Db        159 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSV 211

Qy         91 VNITQGQWMSSNTYTCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPL 140
              | :|   :  : ||||:| |   |  : :   |:|  : |       |   :| || |
Db        212 VTVTSSNF-GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPK 270

Qy        141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPV 198
              | |  : ||::||:|:|:  |:  :   |  :  :   : ::   :  ::|  : |:| |
Db        271 DTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGMEVHNAKTKPREEQFNSTFRVVSVLTV 330

Qy        199 DAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCL 257
                :||: |: |:|:| :   | || ::|:|  |:   |:|| |||| || |    ::|||
Db        331 VHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCL 390

Qy        258 IRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVT 317
              ::|||||:|:|:|  | : ::  :: || |  |  :| ||||||:: |:|| |::||: :
Db        391 VKGFYPSDIAVEWESNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFS 446

Qy        318 CRVVHEALPGSRTLEKSLHYSAG 340
              | |:||||    | :|||  | |
Db        447 CSVMHEALHNHYT-QKSLSLSPG 468


RESULT 12
US-08-437-642B-23
```

```
; Sequence 23, Application US/08437642B
; Patent No. 6054297
;  GENERAL INFORMATION:
;    APPLICANT:  Paul J. Carter
;    APPLICANT:  Leonard G. Presta
;    TITLE OF INVENTION:  Immunoglobulin Variants
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/437,642B
;      FILING DATE:  09-May-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/934373
;      FILING DATE:  21-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/146206
;      FILING DATE:  17-NOV-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/05126
;      FILING DATE:  15-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/715272
;      FILING DATE:  14-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  40,378
;      REFERENCE/DOCKET NUMBER:  P0709P2C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-1994
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  469 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-437-642B-23

  Query Match             25.6%;  Score 475;  DB 3;  Length 469;
  Best Local Similarity   34.7%;  Pred. No. 6.2e-39;
  Matches  112;  Conservative   66;  Mismatches  115;  Indels   30;  Gaps   12;

Qy          35 RGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSE 90
               |  : ||  | |||  : |  | |:|     |  |    : ||       :    :|| |
Db         159 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSV 211

Qy          91 VNITQGQWMSSNTYTCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPL 140
               | :|   :  : ||||:| |   |  : :   |:|  : |         |   :| || |
Db         212 VTVTSSNF-GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPK 270

Qy         141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPV 198
               | |  : ||::||:|:|:  |:  :   |  :  :   : ::   :  ::|  : |:| |
Db         271 DTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGMEVHNAKTKPREEQFNSTFRVVSVLTV 330

Qy         199 DAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCL 257
                 :||: |: |:|:|  :   | || ::|:|  |:   |:|| |||| || |    ::|||
Db         331 VHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCL 390

Qy         258 IRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVT 317
               ::|||||:|:|:|  | : ::  :: || |  |  :| ||||||:: |:|| |::||: :
Db         391 VKGFYPSDIAVEWESNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFS 446

Qy         318 CRVVHEALPGSRTLEKSLHYSAG 340
               | |:||||    | :|||  | |
Db         447 CSVMHEALHNHYT-QKSLSLSPG 468


RESULT 13
US-08-146-206C-23
; Sequence 23, Application US/08146206C
; Patent No. 6407213
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, Paul J.
;    APPLICANT:  Presta, Leonard G.
;    TITLE OF INVENTION:  Method for Making Humanized Antibodies
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/146,206C
;      FILING DATE:  17-No. 6407213-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/715272
;      FILING DATE:  14-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  40,378
;      REFERENCE/DOCKET NUMBER:  P0709P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-1994
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  469 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-146-206C-23

  Query Match             25.6%;  Score 475;  DB 4;  Length 469;
  Best Local Similarity   34.7%;  Pred. No. 6.2e-39;
  Matches  112;  Conservative   66;  Mismatches  115;  Indels   30;  Gaps   12;

Qy          35 RGDAHSTIQLLCLVSGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSE 90
               |  : ||  | |||  : |  | |:|     |  |    : ||       :    :|| |
Db         159 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSV 211

Qy          91 VNITQGQWMSSNTYTCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPL 140
               | :|   :  : ||||:| |   |  : :   |:|  : |         |   :| || |
Db         212 VTVTSSNF-GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPK 270

Qy         141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPV 198
               | |  : ||::||:|:|:  |:  :   |  :  :   : ::   :  ::|  : |:| |
Db         271 DTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGMEVHNAKTKPREEQFNSTFRVVSVLTV 330

Qy         199 DAKDWIEGEGYQCRVDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCL 257
                 :||: |: |:|:|  :   | || ::|:|  |:   |:|| |||| || |    ::|||
Db         331 VHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCL 390

Qy         258 IRGFYPSEISVQWLPNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVT 317
               ::|||||:|:|:|  | : ::  :: || |  |  :| ||||||:: |:|| |::||: :
Db         391 VKGFYPSDIAVEWESNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFS 446

Qy         318 CRVVHEALPGSRTLEKSLHYSAG 340
               | |:||||    | :|||  | |
```

```
Db        447 CSVMHEALHNHYT-QKSLSLSPG 468


RESULT 14
US-08-477-460B-4
; Sequence 4, Application US/08477460B
; Patent No. 6034223
;  GENERAL INFORMATION:
;    APPLICANT:  Progenics Pharmaceuticals, Inc.
;    TITLE OF INVENTION:  NON-PEPTIDYL MOIETY-CONJUGATED
;    TITLE OF INVENTION:  CD4-GAMMA2 AND CD4-IgG2 IMMUNOCONJUGATES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,460B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/927,931
;      FILING DATE:  07-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  41215-A-PCT/JPW/AJM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 977-9550
;      TELEFAX:  (212) 977-9809
;      TELEX:  422523 COOP UI
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  530 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  homo sapien
;      CELL TYPE:  lymphocyte
US-08-477-460B-4

  Query Match             25.3%;  Score 469.5;  DB 3;  Length 530;
  Best Local Similarity   32.8%;  Pred. No. 2.6e-38;
  Matches  121;  Conservative   68;  Mismatches  129;  Indels   51;  Gaps   14;

Qy         10 LSLPESGPVTI------------IPPTVKLFHSSCDP---------RGDAHSTIQLLCLV 48
              | | :|| |              |   | | |:  |          | : ||  | |||
Db        174 LELQDSGTWTCTVLQNQKKVEFKIDIVVLAFASTKGPSVFPLAPCSRSTSESTAALGCLV 233

Qy         49 SGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTY 104
                : |  | |:|    |  |    : ||       :    :|| | | :    :  : ||
Db        234 KDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSVVTVPSSNF-GTQTY 285

Qy        105 TCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPLD-LYENGTPKLTCL 153
              ||:| |   |  : :   |:|  : |         |   :| || | | |  : ||::||:
Db        286 TCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTLMISRTPEVTCV 345

Qy        154 VLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCR 212
              |:|:  |:  :   |  :  :   : ::   :  ::|  : |:| |  :||: |: |:|:
Db        346 VVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCK 405

Qy        213 VDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCLIRGFYPSEISVQWL 271
              | :   | || ::|:|  |:   |:|| |||| || |    ::|||::|||||:|:|:|
Db        406 VSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWE 465

Qy        272 PNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTL 331
              | : ::  :: || |  |  :| ||||||:: |:|| |::||: :| |:||||    |
Db        466 SNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT- 520

Qy        332 EKSLHYSAG 340
              :|||  | |
Db        521 QKSLSLSPG 529


RESULT 15
US-08-379-516-4
; Sequence 4, Application US/08379516
; Patent No. 6083478
; GENERAL INFORMATION:
;  APPLICANT: Allaway, Graham P.
;  APPLICANT: Maddon, Paul J.
;  TITLE OF INVENTION: No. 6083478-Peptidyl Moiety-Conjugated CD4-Gamma2 and CD4-IgG2
;  TITLE OF INVENTION: Immunoconjugates and Uses Thereof
;  FILE REFERENCE: 41215-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/08/379,516
;  CURRENT FILING DATE: 1996-06-10
;  EARLIER APPLICATION NUMBER: PCT/US93/07422
;  EARLIER FILING DATE: 1993-08-06
;  EARLIER APPLICATION NUMBER: 07/927,931
;  EARLIER FILING DATE: 1992-08-07
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-379-516-4

  Query Match             25.3%;  Score 469.5;  DB 3;  Length 530;
  Best Local Similarity   32.8%;  Pred. No. 2.6e-38;
  Matches  121;  Conservative   68;  Mismatches  129;  Indels   51;  Gaps   14;

Qy         10 LSLPESGPVTI------------IPPTVKLFHSSCDP---------RGDAHSTIQLLCLV 48
              | | :|| |              |   | | |:  |          | : ||  | |||
Db        174 LELQDSGTWTCTVLQNQKKVEFKIDIVVLAFASTKGPSVFPLAPCSRSTSESTAALGCLV 233

Qy         49 SGFSPAKVHVTW----LVDGQEAENLFPYTTRPKREGGQTFSLQSEVNITQGQWMSSNTY 104
                : |  | |:|    |  |    : ||       :    :|| | | :    :  : ||
Db        234 KDYFPEPVTVSWNSGALTSG---VHTFPAVL----QSSGLYSLSSVVTVPSSNF-GTQTY 285

Qy        105 TCHVKH---NGSIFEDSSRRCSDDEP-------RGVITYLIPPSPLD-LYENGTPKLTCL 153
              ||:| |   |  : :   |:|  : |         |   :| || | | |  : ||::||:
Db        286 TCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTLMISRTPEVTCV 345

Qy        154 VLDLESEE-NITVTWVRERKKSIGSASQRSTKHHHATTSITSILPVDAKDWIEGEGYQCR 212
              |:|:  |:  :   |  :  :   : ::   :  ::|  : |:| |  :||: |: |:|:
Db        346 VVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCK 405

Qy        213 VDHPHFPKPIVRSITKLPGKRLAPEVYMLPPSPEE-TGTTRTVTCLIRGFYPSEISVQWL 271
              | :   | || ::|:|  |:   |:|| |||| || |    ::|||::|||||:|:|:|
Db        406 VSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWE 465

Qy        272 PNNEEDHTGHHTTTRPQKDHGTDPSFFLYSRMLVNKSIWEKGNLVTCRVVHEALPGSRTL 331
              | : ::  :: || |  |  :| ||||||:: |:|| |::||: :| |:||||    |
Db        466 SNGQPEN--NYKTTPPMLD--SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT- 520

Qy        332 EKSLHYSAG 340
              :|||  | |
Db        521 QKSLSLSPG 529


Search completed: July  9, 2003, 13:40:23
Job time : 17 secs
```